

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 : Search time 1.12756 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNKGAIIGLMVGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	100.0	57	2 E60045	Alzheimer's diseases
2	222	100.0	57	2 F60045	Alzheimer's diseases
3	222	100.0	57	2 D60045	Alzheimer's diseases
4	222	100.0	57	2 D60045	Alzheimer's diseases
5	222	100.0	57	2 A60045	Alzheimer's diseases
6	222	100.0	57	2 B60045	Alzheimer's diseases
7	222	100.0	82	2 P00438	Alzheimer's diseases
8	222	100.0	695	1 A49795	Alzheimer's diseases
9	222	100.0	770	1 QRH044	Alzheimer's diseases
10	217	97.7	42	2 P80512	beta-amyloid prote
11	203	91.4	695	2 A27485	Alzheimer's diseases
12	203	91.4	695	2 S00550	Alzheimer's diseases
13	203	91.4	747	2 JH0773	Alzheimer's diseases
14	133	58.9	33	2 S23094	beta-amyloid prote
15	64	28.8	755	2 A13228	tryptophan 2-monoo
16	62	27.9	755	1 DAAGWT	tryptophan 2-monoo
17	62	27.9	755	1 OQAG4T	genome polyprotein
18	57	25.7	327	2 S11435	general amino acid
19	57	25.7	503	2 S73843	tryptophan - yea
20	56.5	25.5	378	2 S61992	formylmethanofuran
21	55.5	25.0	297	2 G59525	Na+/H+ antiporter
22	55.5	25.0	621	2 AF3016	probable sodium/hy
23	55.5	25.0	642	2 B98268	3-methyl-2-oxobuta
24	55.5	25.0	678	2 G71526	glycosyl transfera
25	55	24.8	291	2 F95015	probable aldehyde
26	55	24.8	317	2 H97888	conserved hypothet
27	55	24.8	488	2 S27652	genome polyprotein
28	55	24.8	738	2 C95936	
29	55	24.8	3063	2 J50166	

30 54.5 24.5 678 2 C81683
31 54.5 24.5 832 2 H84848
32 54 24.3 77 2 C97027
33 54 24.3 284 2 S04723
34 54 24.3 316 2 D97865
35 54 24.3 322 2 E71647
36 54 24.3 763 2 A13443
37 53.5 24.1 245 2 A61461
38 53.5 24.1 245 2 A61098
39 53.5 24.1 708 2 T24727
40 53.5 24.1 971 2 D70128
41 53 23.9 256 2 G96774
42 53 23.9 390 2 G75103
43 53 23.9 422 2 D72302
44 53 23.9 601 2 T02581
45 53 23.9 1555 2 J70959

ALIGNMENTS

RESULT 1

E60045

Alzheimer's disease amyloid beta/A4 protein precursor - sheep (fragment)

C:Species: Ovis sp. (sheep)

C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995

C:Accession: E60045

R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d

A:Reference number: A60045; MUID:92017079; PMID:1656157

A:Accession: E60045

A:Molecule type: mRNA

A:Residues: 1-57 <JOH>

A:Cross-references: EMBL:X56130

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type protein

C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.9e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 48

RESULT 2

F60045

Alzheimer's disease amyloid beta/A4 protein precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999

C:Accession: F60045

R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.

Brain Res. Mol. Brain Res. 10, 299-305, 1991

A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d

A:Reference number: A60045; MUID:92017079; PMID:1656157

A:Accession: F60045

A:Molecule type: mRNA

A:Residues: 1-57 <JOH>

A:Cross-references: EMBL:X56127; NID:q1895; PIDN:CAA39592.1; PTD:g1896

C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type protein

C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;

Best Local Similarity 100.0%; Pred. No. 2.9e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 43

Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGVVIAT 48

RESULT 3
G60045
Alzheimer's disease amyloid beta/A4 protein precursor - guinea pig (fragment)
C:Species: Cavia porcellus (guinea pig)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: G60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079; PMID:1656157
A:Accession: G60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56126
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 4
D60045
Alzheimer's disease amyloid beta/A4 protein precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: D60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079; PMID:1656157
A:Accession: D60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56124
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 5
A60045
Alzheimer's disease amyloid beta/A4 protein precursor - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 28-Jul-1995
C:Accession: A60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog,
A:Reference number: A60045; MUID:92017079; PMID:1656157
A:Accession: A60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56125
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 6
B60045
Alzheimer's disease amyloid beta/A4 protein precursor - polar bear (fragment)
C:Species: Ursus maritimus (polar bear)
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999
C:Accession: B60045
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079; PMID:1656157
A:Accession: B60045
A:Molecule type: mRNA
A:Residues: 1-57 <JOH>
A:Cross-references: EMBL:X56128; NID:g2165; PIDN:CAA39593.1; PID:g2166
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; brain

Query Match 100.0%; Score 222; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.9e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 7
PQ0438
Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C:Accession: PQ0438; C60045
R:Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.
Biochem. Biophys. Res. Commun. 188, 905-911, 1992
A:Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precurs
A:Reference number: PQ0438; MUID:93075180; PMID:1445331
A:Accession: PQ0438
A:Molecule type: DNA
A:Residues: 1-82 <DAV>
A:Cross-references: GB:M83558; GB:M83657
R:Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A:Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in d
A:Reference number: A60045; MUID:92017079; PMID:1656157
A:Accession: C60045
A:Molecule type: mRNA
A:Residues: 12-68 <JOH>
A:Cross-references: EMBL:X56129
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

Query Match 100.0%; Score 222; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 4.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 17 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 59

RESULT 8
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A49795
R:Podlisky, M.B.; Tolan, D.R.; Selkoe, D.J.
Am. J. Pathol. 138, 1423-1435, 1991

A:Title: Homology of the amyloid beta protein precursor in monkey and human supports a P
A:Reference number: A49795; MUID:91273117; PMID:1905108
A:Accession: A49795
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-695 <POD>
A:CROSS-references: GB:M58727; NID:q342062; PIDN:AAA36829.1; PID:q342063
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase i
C:Keywords: alternative splicing

Query Match 100.0%; Score 222; DB 1; Length 695;
Best Local Similarity 100.0%; Pred. No. 4.5e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEYHOKLVFFAEVDGSKNGAIIIGLMVGWVAT 43
|||||
Db 597 DAEFRHDSGYEYHOKLVFFAEVDGSKNGAIIIGLMVGWVAT 639

RESULT 9
ORHU44
Alzheimer's disease amyloid beta protein precursor [validated] - human
N:Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor X1a inhibi
N:Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular
protein precursor splice form APP(770)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence-revision 28-Jul-1995 #text-change 15-Sep-2000
A:Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44
4668; A29583; A29302; A60805; J10038; S06121; A50335; A59011; A38384; S29076; S38252; S3
R:Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayne, R.M.; Unterbeck, A.; Bey
Nucleic Acids Res. 17, 517-522, 1989
A:Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded b
A:Reference number: S02260; MUID:89128427; PMID:2763775
A:Accession: S02260
A:Molecule type: DNA
A:Residues: 1-288, 'V', 365-770 <LEMI>
A:CROSS-references: EMBL:X13466
A:Note: alternative splice form APP(695)
R:Lemaire, H.G.
submitted to the EMBL Data Library, November 1988
A:Reference number: S05194
A:Accession: S05194
A:Molecule type: DNA
A:Residues: 1-14, 'W', 17-288, 'V', 365-770 <LEW2>
A:CROSS-references: EMBL:X13466; NID:q35598; PIDN:CAA31830.1; PID:q871360
A:Note: alternative splice form APP(695)
R:La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
Biochem. Biophys. Res. Commun. 159, 297-304, 1989
A:Title: Characterization of the 5'-end region and the first two exons of the beta-prote
A:Reference number: A32277; MUID:89165870; PMID:2538123
A:Accession: A32277
A:Molecule type: DNA
A:Residues: 1-75 <LAF>
A:CROSS-references: GB:M24546; GB:M24547; NID:q341202; PIDN:AA013654.1; PID:q516074
R:Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
A:Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarit
A:Reference number: A33260; MUID:89392030; PMID:2675837
A:Accession: A33260
A:Molecule type: DNA
A:Residues: 656-737 <JOH>
A:CROSS-references: NID:q178863; PIDN:AAA51768.1; PID:q178865
R:Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A:Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of
A:Reference number: A35486; MUID:90321244; PMID:2196878
A:Accession: A35486
A:Molecule type: DNA
A:Residues: 672-710 <PRE1>
A:Note: 693-Gln was found in DNA isolated from HCHWA-D patients
R:Yoshikai, S.I.; Sakaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 87, 257-263, 1990
A:Title: Genomic organization of the human amyloid beta-protein precursor gene.

A:Reference number: I39451; MUID:90236318; PMID:2110105
A:Accession: I39452
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-770 <YOS1>
A:CROSS-references: GB:M33112; NID:q178613; PIDN:AAB59502.1; PID:q178616
A:Accession: I39451
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-530, 'OWLMPIVPAFWFAKVGK' <YOS2>
A:CROSS-references: GB:M34875; NID:q178608; PIDN:AAB59501.1; PID:q178615
R:Yoshikai, S.I.; Sakaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A:Reference number: A59020; MUID:91340168; PMID:1908403
A:Contents: annotation; erratum
A:Note: revised physical map for reference I39451
R:Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Du
Science 248, 1124-1126, 1990
A:Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemo
A:Reference number: I39453; MUID:90260663; PMID:2111384
A:Accession: I39453
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 656-737 <LEV>
A:CROSS-references: GB:M37896; NID:q178618; PIDN:AAA51727.1; PID:q178620
A:Note: a mutation with 693-Gln is presented
R:Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A:Title: A mutation in the amyloid precursor protein associated with hereditary Alzhe
A:Reference number: I59562; MUID:92022553; PMID:1925564
A:Accession: I59562
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 689-716, 'F', 718-737 <MUR>
A:CROSS-references: GB:S57665; NID:q236720; PIDN:AAB19991.1; PID:q236721
R:Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anders
arakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Mart
Am. J. Hum. Genet. 51, 998-1014, 1992
A:Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for t
A:Reference number: A44017; MUID:93035397; PMID:1415269
A:Accession: A44017
A:Molecule type: DNA
A:Residues: 687-692, 'G', 694-718 <KAM1>
A:CROSS-references: GB:S45135; NID:q257377; PIDN:AAB23645.1; PID:q257378
A:Experimental source: familial Alzheimer disease family SB
A:Note: sequence extracted from NCBI backbone (NCBIP:115374)
A:Accession: B44017
A:Molecule type: DNA
A:Residues: 687-718 <KAM2>
A:CROSS-references: GB:S45136; NID:q257379; PIDN:AAB23646.1; PID:q257380
A:Experimental source: familial Alzheimer disease family LIT
A:Note: sequence extracted from NCBI backbone (NCBIP:115376)
A:Note: this sequence has a silent mutation
R:Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.
Nature 325, 733-736, 1987
A:Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-sur
A:Reference number: A03134; MUID:87144572; PMID:2881207
A:Accession: A03134
A:Molecule type: mRNA
A:Residues: 1-288, 'V', 365-770 <KAN>
A:CROSS-references: GB:Y00264; NID:q28525; PIDN:CAA68374.1; PID:q28526
A:Note: alternative splice form APP(695)
R:Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A:Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascula
A:Reference number: A29030; MUID:87231971; PMID:3035574
A:Accession: A29030
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-646, 'E', 648-770 <ROB>
A:CROSS-references: GB:M16765; NID:q178539; PIDN:AAA51722.1; PID:q178540
A:Note: the authors translated the codon GAG for residue 647 as Asp
R:Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffioti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987

A:Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid
 A:Reference number: A47584; MUID:87120328; PMID:3810169
 A:Accession: A47584
 A:Molecule type: mRNA
 A:Residues: 674-756, 'S', 758-770 <GOL>
 A:Cross-references: GB:MJ5533; NID:g178706; PIDN:AAA35540.1; PID:g178707
 A:Experimental source: brain
 R:Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Ke
 Science 235, 880-884, 1987
 A:Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near th
 A:Reference number: A47585; MUID:87120329; PMID:2949367
 A:Accession: A47585
 A:Molecule type: mRNA
 A:Residues: 674-703 <TANI>
 A:Cross-references: GB:MJ5533; NID:g177957; PIDN:AAA51564.1; PID:g177958
 R:Dytrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Muelle
 EMBO J. 7, 949-957, 1988
 A:Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 prec
 A:Reference number: S02638; MUID:88296437; PMID:2900137
 A:Accession: S02638
 A:Molecule type: mRNA
 A:Residues: 672-678 <DYR>
 R:Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve
 Nature 331, 528-530, 1988
 A:Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associat
 A:Reference number: S00707; MUID:88122640; PMID:2893290
 A:Accession: S00707
 A:Molecule type: mRNA
 A:Residues: 286-344, 'I', 365-366 <TAN2>
 A:Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612
 A:Experimental source: promyelocytic leukemia cell line HL60
 A:Note: alternative splice form APP(751)
 R:Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.; Da
 Nature 331, 525-527, 1988
 A:Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibit
 A:Reference number: S00925; MUID:88122639; PMID:2893289
 A:Accession: S00925
 A:Molecule type: mRNA
 A:Residues: 1-344, 'I', 365-770 <PO2>
 A:Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
 A:Note: alternative splice form APP(751)
 R:Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
 Nature 331, 530-532, 1988
 A:Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibit
 A:Reference number: A38949; MUID:88122641; PMID:2893291
 A:Accession: A38949
 A:Molecule type: mRNA
 A:Residues: 287-367 <KIT>
 A:Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611
 A:Experimental source: glioblastoma cell line
 A:Note: alternative splice form APP(770)
 R:Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ashton
 Brain Res. Mol. Brain Res. 4, 121-131, 1988
 A:Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three P
 A:Reference number: A30320
 A:Accession: A30320
 A:Molecule type: mRNA
 A:Status: not compared with conceptual translation
 A:Residues: 284-288, 'V', 365-770 <VIT1>
 A:Accession: B03020
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 122-288, 'V', 365-770 <VIT2>
 A:Accession: C03020
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 606-770 <VIT3>
 R:Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta, C.A
 Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
 A:Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease br
 A:Reference number: A31087; MUID:88124954; PMID:2893379
 A:Accession: A31087
 A:Molecule type: mRNA

A:Residues: 507-770 <ZAI>
 A:Cross-references: GB:M18734; NID:g178572; PIDN:AAA51736.1; PID:g178573
 A:Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue
 8 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue
 A:Note: the cited Genbank accession number, J03594, is not in release 101.0
 R:Masters, C.L.; Multhaup, G.; Simms, G.; Potgiesser, J.; Martins, R.N.; Beyreuther,
 Query Match 100.0%; Score 222; DB 1; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5e-21;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 43
 Db 672 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 714
 RESULT 10
 PN0512
 beta-amyloid protein - guinea pig (fragment)
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
 C:Accession: PN0512
 R:Shimohigashi, Y.; Matsumoto, H.; Takano, Y.; Saito, R.; Iwata, T.; Kamiya, H.; Ohno
 Biochem. Biophys. Res. Commun. 193, 624-630, 1993
 A:Title: Receptor-mediated specific biological activity of a beta-amyloid protein fra
 A:Reference number: PN0512; MUID:93290653; PMID:7685598
 A:Accession: PN0512
 A:Molecule type: protein
 A:Residues: 1-42 <SHT>
 C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinas
 C:Keywords: alternative splicing; amyloid
 Query Match 97.7%; Score 217; DB 2; Length 42;
 Best Local Similarity 100.0%; Pred. No. 9.3e-22;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
 Db 1 DAERHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
 RESULT 11
 A27485
 Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
 N:Alternate names: proteinase nexin II
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
 C:Accession: A27485; S19727; I49485
 R:Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sasaki, Y.
 Biochem. Biophys. Res. Commun. 149, 665-671, 1987
 A:Title: Complementary DNA for the mouse homolog of the human amyloid beta protein pr
 A:Reference number: A27485; MUID:88106489; PMID:3322280
 A:Accession: A27485
 A:Molecule type: mRNA
 A:Residues: 1-695 <YAM>
 A:Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
 A:Experimental source: brain
 R:de Strooper, B.; van Leuven, F.; van den Berghe, H.
 Biochim. Biophys. Acta 1129, 141-143, 1991
 A:Title: The amyloid beta protein precursor or proteinase nexin II from mouse is clos
 A:Reference number: S19727; MUID:92096458; PMID:1756177
 A:Accession: S19727
 A:Molecule type: mRNA
 A:Residues: 1-210, 'G', 212-220, 'S', 222-396, 'A', 398-402, 'T', 404-448, 'A', 450-695 <STR>
 A:Cross-references: EMBL:X59379
 R:Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sasaki, Y.
 Gene 112, 189-195, 1992
 A:Title: Positive and negative regulatory elements for the expression of the Alzheimer
 A:Reference number: I49485; MUID:92209998; PMID:1555768
 A:Accession: I49485
 A:Status: translated from GB/EMBL/DBDJB
 A:Molecule type: DNA
 A:Residues: 1-19 <RES>

Matches 16; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

OY 7 DSGYEVHHOKLVFFAEDVGSNKGAIIGLWVGWVIAT 43

||| :||| | ||| :|:|:

Db 223 DSG-----RIGFFPEDVPKPKVAIIGAGISGLVVAS 253

Search completed: January 28, 2003, 14:19:54
Job time : 4.12756 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 0.657744 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNGKAIIGLMVGGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	222	100.0	57	A4_PIG	Q29023 sus scrofa
2	222	100.0	57	A4_URSMA	Q29149 ursus marit
3	222	100.0	58	A4_CANFA	Q28280 canis famil
4	222	100.0	58	A4_RABIT	Q28748 oryctolagus
5	222	100.0	58	A4_SHEEP	Q28757 ovis aries
6	222	100.0	59	A4_BOVIN	Q28053 bos taurus
7	222	100.0	751	A4_SALISC	Q95341 salmtr sci
8	222	100.0	770	A4_HUMAN	P05067 homo sapien
9	203	91.4	770	A4_MOUSE	P12023 mus musculu
10	203	91.4	770	A4_RAT	P12023 mus musculu
11	62	27.9	755	TR2N_AGRV1	P08592 rattus norv
12	62	27.9	755	TR2N_AGRV1	P04029 agrobacteri
13	57	25.7	327	POLG_PVYCH	P25017 agrobacteri
14	57	25.7	503	Y226_MYCPN	P21294 potato viru
15	56.5	25.5	378	SLGI_YEAST	P75462 mycoplasma
16	55.5	25.0	297	FTN_ARCFU	P54867 saccharomyc
17	55	24.8	488	HAL_PSESP	Q28076 archaeoglob
18	55	24.8	3063	POLG_PVYN	P33008 pseudomonas
19	54.5	24.5	967	PDB1_ARATH	P18247 p genome po
20	54	24.3	284	POLG_PVYIO	P93733 arabidopsis
21	54	24.3	322	Y853_RICPR	P11897 potato viru
22	53.5	24.1	708	YN2B_CAEEL	Q92ca7 rickettsia
23	53.5	24.1	971	Y228_BORBU	P45972 caenorhabdi
24	52	23.4	611	YCR3_YEAST	O51246 borrelia bu
25	51.5	23.2	1036	PLD1_CRIGR	P25351 saccharomyc
26	51	23.0	494	COBQ_MYCTU	O08684 cricetulus
27	51	23.0	763	APP2_HUMAN	O53677 mycobacteri
28	51	23.0	769	ITB2_BOVIN	Q06481 homo sapien
29	51	23.0	3579	STAN_DROME	P23592 bos taurus
30	50.5	22.7	915	PDB2_ARATH	Q9v5n8 drosophila
31	50	22.5	285	1	Q23078 arabidopsis
32	50	22.5	674	1	P44960 haemophilus
33	50	22.5	1437	1	P33543 arabidopsis
34	50	22.5	1437	1	O15440 homo sapien

34 49.5 22.3 1162 1 VGL2_IBVM
35 49 22.1 246 1 TPIS_CULTA
36 49 22.1 403 1 PGK_STRCO
37 49 22.1 432 1 PURA_RHILO
38 49 22.1 542 1 YM87_MYCTU
39 49 22.1 568 1 NIRS_PSEAE
40 49 22.1 695 1 APP2_MOUSE
41 49 22.1 701 1 YGSL_YEAST
42 49 22.1 704 1 SSP2_BOMMO
43 49 22.1 753 1 PPE2_HUMAN
44 49 22.1 757 1 PPE2_MOUSE
45 49 22.1 1080 1 HDC_DROME

ALIGNMENTS

RESULT 1

A4_PIG STANDARD; PRT; 57 AA.
AC Q29023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog (Contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)) (Fragment).
GN APP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA "Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis."
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC HMBP; X56127; CAA39592.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
DR PROSITE; PS00319; A4_EXTRA; PARTIAL.
DR PROSITE; PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein; Amyloid; Neutrone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT NON_TER 57 57
SQ SEQUENCE 57 AA; 6172 MW; 84209D88EBA82DFA CRC64;

Query Match 100.0%; Score 222; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 2.7e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 ID A4_URSWA STANDARD: PRT; 57 AA.
 Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 2

AC Q29149;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
 DE protein (Beta-APP) (A-beta)] (Fragment).
 GN APP.
 OS Ursus maritimus (Polar bear) (Thalarcos maritimus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.
 OX NCBI_TaxID=29073;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92017079; PubMed=1656157;
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
 RT "Conservation of the sequence of the Alzheimer's disease amyloid
 RT peptide in dog, polar bear and five other mammals by cross-species
 RT polymerase chain reaction analysis.";
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).
 CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X56128; CAA39593.1; -;
 CC HSP; P05067; IBA4.
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR001255; Beta-APP.
 CC Pfam; PF03494; Beta-APP; 1.
 CC PROSITE; PS00319; A4_EXTRA; PARTIAL.
 CC PROSITE; PS00320; A4_INTRA; PARTIAL.
 CC Glycoprotein; Amyloid; Neurone; Transmembrane.
 CC NON_TER 1 1
 CC CHAIN 1 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
 CC DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 34 57 POTENTIAL.
 CC NON_TER 57 57
 CC SEQUENCE 57 AA; 6172 MW; 8420988EBA82DFA CRC64;

Query Match 100.0%; Score 222; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 2.7e-22;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 ID A4_CANFA STANDARD: PRT; 58 AA.
 Db 6 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 48

RESULT 3

AC Q28280;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
 DE protein (Beta-APP) (A-beta)] (Fragment).
 GN APP.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=92017079; PubMed=1656157;
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
 RT "Conservation of the sequence of the Alzheimer's disease amyloid
 RT peptide in dog, polar bear and five other mammals by cross-species
 RT polymerase chain reaction analysis.";
 RL Brain Res. Mol. Brain Res. 10:299-305(1991).
 CC -1- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X56125; CAA39590.1; -;
 CC HSP; P05067; IBA4.
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR001255; Beta-APP.
 CC Pfam; PF03494; Beta-APP; 1.
 CC PROSITE; PS00319; A4_EXTRA; PARTIAL.
 CC PROSITE; PS00320; A4_INTRA; PARTIAL.
 CC Glycoprotein; Amyloid; Neurone; Transmembrane.
 CC NON_TER 1 1
 CC CHAIN 1 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).
 CC DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 35 58 POTENTIAL.
 CC NON_TER 58 58
 CC SEQUENCE 58 AA; 6285 MW; 8469D488A2E12DFA CRC64;

Query Match 100.0%; Score 222; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.8e-22;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 7 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 49

RESULT 4

A4_RABIT
 ID A4_RABIT STANDARD: PRT; 58 AA.
 AC Q28748;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
 DE protein (Beta-APP) (A-beta)] (Fragment).
 GN APP.

OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain;
 RX MEDLINE=92017079; PubMed=1656157;
 RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;

RT "Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis."

RL Brain Res. Mol. Brain Res. 10:299-305(1991).

CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O) (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X56129; CAA39594.1; -.

DR HSSP; P05067; IBA4.

DR InterPro; IPR001868; A4_APP.

DR InterPro; IPR001255; Beta-APP.

DR Pfam; PF03494; Beta-APP; 1.

DR PROSITE; PS00319; A4_EXTRA; PARTIAL.

DR PROSITE; PS00320; A4_INTRA; PARTIAL.

KW Glycoprotein; Amyloid; Neurone; Transmembrane.

FT NON_TER 1 1

FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).

FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 34 57 POTENTIAL.

FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).

FT NON_TER 58 58

SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;

Query Match 100.0%; Score 222; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHVKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||

Db 6 DAEFRHDSGYEVHVKLVFFAEADVGSNKGAIIGLMVGGVVIAT 48
|||||

RESULT 5

A4_SHEEP

ID A4_SHEEP STANDARD; PRT; 58 AA.

AC Q28757;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid protein (Beta-APP) (A-beta)] (Fragment).

GN APP.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.

OX NCBI_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RX MEDLINE=92017079; PubMed=1656157;

RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
"Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis."

RT Brain Res. Mol. Brain Res. 10:299-305(1991).

CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O) (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X56130; CAA39595.1; -.

DR HSSP; P05067; IBA4.

DR InterPro; IPR001868; A4_APP.

DR InterPro; IPR001255; Beta-APP.

DR Pfam; PF03494; Beta-APP; 1.

DR PROSITE; PS00319; A4_EXTRA; PARTIAL.

DR PROSITE; PS00320; A4_INTRA; PARTIAL.

KW Glycoprotein; Amyloid; Neurone; Transmembrane.

FT NON_TER 1 1

FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).

FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 34 57 POTENTIAL.

FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).

FT NON_TER 58 58

SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;

Query Match 100.0%; Score 222; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHVKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||

Db 6 DAEFRHDSGYEVHVKLVFFAEADVGSNKGAIIGLMVGGVVIAT 48
|||||

RESULT 6

A4_BOVIN

ID A4_BOVIN STANDARD; PRT; 59 AA.

AC Q28053;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid protein (Beta-APP) (A-beta)] (Fragment).

GN APP.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=92017079; PubMed=1656157;

RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
"Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis."

RT Brain Res. Mol. Brain Res. 10:299-305(1991).

CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O) (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X56124; CAA39589.1; -.

DR EMBL; X56126; CAA39591.1; -.


```

DR HSSP; P05067; 1BA4.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
DR PROSITE; PS00319; A4_EXTRA; PARTIAL.
DR PROSITE; PS00320; A4_INTRA; PARTIAL.
DR Glycoprotein; Amyloid; Neurone; Transmembrane.
FT CHAIN 1 49 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 35 58 POTENTIAL.
FT DOMAIN 59 >59 CYTOPLASMIC (POTENTIAL).
FT NON_TER 59 59
SQ SEQUENCE 59 AA; 6414 MW; F43469D488A2E12D CRC64;

Query Match 100.0%; Score 222; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 7 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 49

RESULT 7
A4_SAISC STANDARD; PRT; 751 AA.
ID A4_SAISC
AC Q95241;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein precursor [Contains: Beta-amyloid protein (Beta-APP) (A-beta)].
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Platyrrhini; Cebinae; Cebidae; Saimiri.
ON NCBI_TaxID=9521;
RX MEDLINE=96108492; PubMed=85322114;
SEQUENCE FROM N.A.
TI Tissue=Liver, and Kidney;
RA Levy E., Amorim A., Frangione B., Walker L.C.;
RT "Beta-amyloid precursor protein gene in squirrel monkeys with cerebral amyloid angiopathy."
RL Neurobiol. Aging 16:805-808(1995).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION WITH XI1-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF PHOSPHORYLATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; S81024; AAD14347.1; -.
CC HSSP; P05067; 1AAP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz-BPTI.
DR Pfam; PF00014; Kunitz-BPTI; 1.
DR Pfam; PF02177; A4_EXTRA; 1.

DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz-BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Glycoprotein; Amyloid; Neurone; Transmembrane; Alternative splicing;
FT SIGNAL 1 17 Serine protease inhibitor.
FT CHAIN 18 751 A4 PROTEIN.
FT CHAIN 553 695 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN 18 680 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 681 704 POTENTIAL.
FT DOMAIN 705 751 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
FT SITE 740 743 CLATHRIN-BINDING (BY SIMILARITY).
FT ACT_SITE 301 302 REACTIVE BOND.
FT DISULFID 291 341 BY SIMILARITY.
FT DISULFID 300 324 BY SIMILARITY.
FT DISULFID 316 337 BY SIMILARITY.
FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 552 552 N-LINKED (GLCNAC. . .) (PROBABLE).
SQ SEQUENCE 751 AA; 84893 MW; 6C3E431089569049 CRC64;

Query Match 100.0%; Score 222; DB 1; Length 751;
Best Local Similarity 100.0%; Pred. No. 3.8e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 653 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 695

RESULT 8
A4_HUMAN STANDARD; PRT; 770 AA.
ID A4_HUMAN
AC P05067; P09000; Q16011;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein precursor (Protease nexin-II) (PN-II) (APPI) [Contains: Beta-amyloid protein (Beta-APP) (A-beta)].
GN APP OR A4 OR CVAP OR ADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87144572; PubMed=2881207;
RA Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L., Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
RT "The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor."
RL Nature 325:733-736(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88122639; PubMed=2893289;
RA Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D., Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F., Cordell B.;
RT "A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors."
RL Nature 331:525-527(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89128427; PubMed=2783775;
RA Lemaire H.G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,

```


RA Untebeck A., Beyreuther K., Mueller-Hill B.;
 RT "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
 RL is encoded by 16 exons.";
 RN Nucleic Acids Res. 17:517-522(1989).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97263807; PubMed=9108164;
 RA Hattori M., Tsukahara F., Furuhashi Y., Tanahashi H., Hirose M.,
 RA Saito M., Tsukuni S., Sakaki Y.;
 RT "A novel method for making nested deletions and its application for
 RT sequencing of a 300 kb region of human APP locus.";
 RL Nucleic Acids Res. 25:1802-1808(1997).
 RN [5]
 RP SEQUENCE OF 286-345 AND 365-366 FROM N.A.
 RX MEDLINE=88122640; PubMed=2893290;
 RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
 RA Gusella J.F., Neve R.L.;
 RT "Protease inhibitor domain encoded by an amyloid protein precursor
 RT mRNA associated with Alzheimer's disease.";
 RL Nature 331:528-530(1988).
 RN [6]
 RP SEQUENCE OF 287-367 FROM N.A.
 RX MEDLINE=88122641; PubMed=2893291;
 RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
 RT "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [7]
 RP SEQUENCE OF 284-289 AND 365-770 FROM N.A.
 RX MEDLINE=87231971; PubMed=3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RT cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [8]
 RP SEQUENCE OF 507-770 FROM N.A.
 RX MEDLINE=88124954; PubMed=2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RT disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [9]
 RP SEQUENCE OF 672-681.
 RX MEDLINE=88035004; PubMed=3312495;
 RA Partridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtellotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RT and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).
 RN [10]
 RP SEQUENCE OF 739-770 FROM N.A.
 RX MEDLINE=90236318; PubMed=2110105;
 RA Yoshikai S.I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RT gene.";
 RL Gene 87:257-263(1990).
 RN [11]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE=89016647; PubMed=3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RT encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [12]
 RP SEQUENCE OF 18-50.
 RX MEDLINE=87250462; PubMed=3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 262:8508-8514(1987).

RN [13]
 RP IDENTITY OF APP WITH NEXIN-II.
 RX MEDLINE=89384866; PubMed=2506449;
 RA Oltersdorf T., Fritz L.C., Schenk D.B., Lieberburg I.,
 RA Johnson-Wood K.L., Beattie E.C., Ward P.J., Blacher R.W., Dovey H.F.,
 RA Sinha S.;
 RT "The secreted form of the Alzheimer's amyloid precursor protein with
 RT the Kunitz domain is protease nexin-II.";
 RL Nature 341:144-147(1989).
 RN [14]
 RP PROTEASE-SPECIFICITY OF INHIBITOR DOMAIN.
 RX MEDLINE=90211252; PubMed=1969731;
 RA Kido H., Fukutomi A., Schilling J., Wang Y., Cordell B., Katunuma N.;
 RT "Protease-specificity of Kunitz inhibitor domain of Alzheimer's
 RT disease amyloid protein precursor.";
 RL Biochem. Biophys. Res. Commun. 167:716-721(1990).
 RN [15]
 RP COMPLEX WITH G(O).
 RX MEDLINE=93188965; PubMed=8446172;
 RA Nishimoto I., Okamoto T., Matsuura Y., Takahashi S., Okamoto T.,
 RA Murayama Y., Ogata E.;
 RT "Alzheimer amyloid protein precursor complexes with brain GTP-binding
 RT protein G(O).";
 RL Nature 362:75-79(1993).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 28-133.
 RX MEDLINE=99215582; PubMed=10201399;
 RA Rossjohn J., Cappai R., Feil S.C., Henry A., McKinstry W.J.,
 RA Galatis D., Hesse L., Multhaup G., Beyreuther K., Masters C.L.,
 RA Parker M.W.;
 RT "Crystal structure of the N-terminal, growth factor-like domain of
 RT Alzheimer amyloid precursor protein.";
 RL Nat. Struct. Biol. 6:327-331(1999).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 287-344.
 RX MEDLINE=91104913; PubMed=2125487;
 RA Hynes T.R., Randal M., Kennedy L.A., Eigenbrot C., Kossiakof A.A.;
 RT "X-ray crystal structure of the protease inhibitor domain of
 RT Alzheimer's amyloid beta-protein precursor.";
 RL Biochemistry 25:10018-10022(1990).
 RN [18]
 RP STRUCTURE BY NMR OF 289-344.
 RX MEDLINE=92031488; PubMed=1718421;
 RA Heald S.L., Tilton R.F. Jr., Hammond L.S., Lee A., Bayney R.M.,
 RA Kamack M.E., Ramabhadran T.V., Dreyer R.N., Davis G., Unterbeck A.,
 RA Tamburini P.P.;
 RT "Sequential NMR resonance assignment and structure determination of
 RT the Kunitz-type inhibitor domain of the Alzheimer's beta-amyloid
 RL Biochemistry 30:10467-10478(1991).
 RN [19]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=94281210; PubMed=7516706;
 RA Talafous J., Marciniowski K.J., Klopman G., Zagorski M.G.;
 RT "Solution structure of residues 1-28 of the amyloid beta-peptide.";
 RL Biochemistry 33:7788-7796(1994).
 RN [20]
 RP STRUCTURE BY NMR OF 696-706.
 RX MEDLINE=97128622; PubMed=8973180;
 RA Kohno T., Kobayashi K., Maeda T., Sato K., Takashima A.;
 RT "Three-dimensional structures of the amyloid beta peptide (25-35) in
 RT membrane-mimicking environment.";
 RL Biochemistry 35:16094-16104(1996).
 RN [21]
 RP STRUCTURE BY NMR OF 672-711.
 RX MEDLINE=98359783; PubMed=9693002;
 RA Coles M., Bicknell W., Watson A.A., Fairlie D.P., Craik D.J.;
 RT "Solution structure of amyloid beta-peptide(1-40) in a water-miscible
 RT environment. Is the membrane-spanning domain where we think it is?";
 RL Biochemistry 37:11064-11077(1998).
 RN [22]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=20400066; PubMed=10940222;

RA Poulsen S.-A., Watson A.A., Craik D.J.;
 RT "Solution structures in aqueous SDS micelles of two amyloid beta
 RT peptides of Abeta(1-28) mutated at the alpha-secretase cleavage
 RT site."
 RL J. Struct. Biol. 130:142-152(2000).
 RN [23]
 RP STRUCTURE BY NMR OF 681-706.
 RX MEDLINE=20400065; PubMed=10940221;
 RA Zhang S., Iwata K., Lachenmann M.J., Peng J.W., Li S., Stimson E.R.,
 RA Lu Y., Felix A.M., Maggio J.E., Lee J.P.;
 RT "The Alzheimer's peptide a beta adopts a collapsed coil structure in
 RT water."
 RL J. Struct. Biol. 130:130-141(2000).
 RN [24]
 RP SIGNAL SEQUENCE CLEAVAGE SITE, AND TOPOLOGY.
 RX MEDLINE=88296437; PubMed=2900137;
 RA Dykx J., Weidemann A., Multhaup G., Salbaum J.M., Lemaire H.-G.,
 RA Kang J., Mueller-Hill B., Masters C.L., Beyreuther K.;
 RT "Identification, transmembrane orientation and biogenesis of the
 RT amyloid A4 precursor of Alzheimer's disease."
 Query Match 100.08; Score 222; DB 1; Length 770;
 Best Local Similarity 100.08; Pred. No. 3.9e-21; Indels 0; Gaps 0;
 Matches 43; Conservative 0; Mismatches 0;
 QY 1 DAEFRHDSGYEVHHQKLVFAEDVGSNGKAIIGLMVGGVVIAT 43
 ID A4_MOUSE STANDARD; PRT; 770 AA.
 AC P12023;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor
 DE (Amyloidogenic glycoprotein) (AG).
 GN APP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC STRAIN=BALB/C; TISSUE=Brain;
 RX MEDLINE=92096458; PubMed=1756177;
 RA de Strooper B., van Leuven F., van den Bergh H.;
 RT "The amyloid beta protein precursor or proteinase nexin II from mouse
 RT is closer related to its human homolog than previously reported."
 RL Biochim. Biophys. Acta 1129:141-143(1991).
 RN [2]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88106489; PubMed=3322280;
 RA Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sasaki Y.;
 RT "Complementary DNA for the mouse homolog of the human amyloid beta
 RT protein precursor."
 RL Biochem. Biophys. Res. Commun. 149:665-671(1987).
 RN [3]
 RP REVISIONS.
 RA Yamada T.;
 RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC STRAIN=CD-1; TISSUE=Placenta;
 RX MEDLINE=89345111; PubMed=2569710;
 RA Fukuchi K., Martin G.M., Deeb S.S.;
 RT "Sequence of the protease inhibitor domain of the A4 amyloid protein
 RT precursor of Mus domestica."
 RL Nucleic Acids Res. 17:5396-5396(1989).
 RN [5]

RP SEQUENCE OF 1-19 FROM N.A.
 RX MEDLINE=9220998; PubMed=1555768;
 RA Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
 RA Sakai Y.;
 RT "Positive and negative regulatory elements for the expression of the
 RT Alzheimer's disease amyloid precursor-encoding gene in mouse."
 RL Gene 112:189-195(1992).
 RN [6]
 RP SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain, and Kidney;
 RX MEDLINE=8919813; PubMed=2493250;
 RA Yamada T., Sasaki H., Dohura K., Goto I., Sasaki Y.;
 RT "Structure and expression of the alternatively-spliced forms of mRNA
 RT for the mouse homolog of Alzheimer's disease amyloid beta protein
 RT precursor."
 RL Biochem. Biophys. Res. Commun. 158:906-912(1989).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(393), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- TISSUE SPECIFICITY: AAA(770) IS EXPRESSED IN KIDNEY. AAA(751) IS
 CC WIDELY EXPRESSED. AAA(695) IS EXPRESSED IN BRAIN, KIDNEY AND
 CC LIVER.
 CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X59379; -; NOT_ANNOTATED_CDS.
 DR EMBL; M18373; AAA37139.1; -;
 DR EMBL; X15210; CAA33280.1; -;
 DR EMBL; D10603; BAA01456.1; -;
 DR EMBL; M24397; AAA33929.1; -;
 DR PIR; A27485; A27485.
 DR PIR; S04855; S04855.
 DR PIR; S19727; S19727.
 DR HSP; P05067; IAA.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
 KW Alternative splicing; Serine protease inhibitor.
 FT SIGNAL 1 17
 FT CHAIN 18 770
 FT ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
 FT HOMOLOG.
 FT DOMAIN 18 699
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 700 723
 FT POTENTIAL.
 FT DOMAIN 724 770
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 673 715
 FT EQUIVALENT OF BETA-AMYLOID PROTEIN.

FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
 FT SITE 759 762 CLATHRIN-BINDING (BY SIMILARITY).
 FT DISULFID 291 341 BY SIMILARITY.
 FT DISULFID 300 324 BY SIMILARITY.
 FT DISULFID 316 337 BY SIMILARITY.
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 289 289 E -> V (IN ISOFORM APP(695)).
 FT VARSPLIC 290 364 MISSING (IN ISOFORM APP(695)).
 FT VARSPLIC 346 380 MISSING (IN ISOFORM APP(751)).
 SQ SEQUENCE 770 AA; 86752 MW; 26C50DE0890CA7A CRC64;

Query Match 91.4%; Score 203; DB 1; Length 770;
 Best Local Similarity 93.0%; Pred. No. 1.1e-18;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAEFRHDSYEVHVKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 43
 DB 672 DAEFGHDSGFVVRHKLFFAEADVGSNKGAIIGLMVGWGVVIAT 714

RESULT 10
 A4_RAT
 ID A4_RAT STANDARD; PRT; 770 AA.
 AC P08592;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein homolog precursor
 DE (Amyloidogenic glycoprotein) (AG).
 GN APP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE OF 1-289 AND 365-770 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88312583; PubMed=2900758;
 RA Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
 RA Seeburg P.H.;
 RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
 in rat brain suggests a role in cell contact.";
 RT EMBO J. 7:1365-1370(1988).
 RN [2]
 RP SEQUENCE OF 289-364 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89183625; PubMed=2648331;
 RA Kang J., Mueller-Hill B.;
 RT "The sequence of the two extra exons in rat preA4.";
 RL Nucleic Acids Res. 17:2130-2130(1989).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
 CC WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
 CC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
 CC NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07648; CAA30488.1; -;
 DR EMBL; X14066; CAA32229.1; -;

DR PIR; S00550; S00550.
 DR PIR; S03607; S03607.
 DR HSP; P05067; IAAP.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF00014; Kunitz_Bpti; 1.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00759; BASICTPASE.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRODOM; PD000222; Kunitz_Bpti; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
 KW Alternative splicing; Serine protease inhibitor.
 FT SIGNAL 1 17
 FT CHAIN 18 770
 FT DOMAIN 18 699
 FT TRANSMEM 700 723
 FT DOMAIN 724 770
 FT DOMAIN 673 715
 FT DOMAIN 287 345
 FT SITE 759 762
 FT DISULFID 291 341
 FT DISULFID 300 324
 FT DISULFID 316 337
 FT CARBOHYD 542 542
 FT CARBOHYD 571 571
 FT VARSPLIC 289 289
 FT VARSPLIC 290 364
 SQ SEQUENCE 770 AA; 86704 MW; C26C9D6BB2D929A7 CRC64;

Query Match 91.4%; Score 203; DB 1; Length 770;
 Best Local Similarity 93.0%; Pred. No. 1.1e-18;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 DAEFRHDSYEVHVKLVFFAEADVGSNKGAIIGLMVGWGVVIAT 43
 DB 672 DAEFGHDSGFVVRHKLFFAEADVGSNKGAIIGLMVGWGVVIAT 714

RESULT 11
 TR2M_AGR4
 ID TR2M_AGR4 STANDARD; PRT; 755 AA.
 AC P04029;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
 GN TWS1.
 OS Agrobacterium tumefaciens (strain Ach5), and
 OS Agrobacterium tumefaciens.
 OG Plasmid pTiAch5, and plasmid pTiA6NC.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176298, 358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ach5; PLASMID=pTiAch5;
 RX MEDLINE=84207942; PubMed=6327292;
 RA Gielen J., de Beuckeleer M., Seurinck J., Deboeck F., de Greve H.,
 RA Lemmers M., van Montagu M., Schell J.;
 RT "The complete nucleotide sequence of the TL-DNA of the Agrobacterium
 tumefaciens plasmid pTiAch5.";
 RL EMBO J. 3:835-846(1984).
 RN [2]
 RP SEQUENCE FROM N.A.


```

RC PLASMID=PTIA6NC;
RX MEDLINE=84170374; PubMed=6584906;
RA Klee H., Montoya A., Horodyski F., Lichtenstein C., Garfinkel D.,
RA Fuller S., Flores C., Peschon J., Nester E., Gordon M.;
RT "Nucleotide sequence of the tms genes of the pTIA6NC octopine Ti
RT plasmid; two gene products involved in plant tumorigenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:1728-1732(1984).
CC
CC -!- CATALYTIC ACTIVITY: L-tryptophan + O(2) = indole-3-acetamide +
CC CO(2) + H(2)O.
CC
CC -!- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF AUXINS FROM TRYPTOPHAN.
CC
CC -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF ACHS.
CC
CC -!- SIMILARITY: SIGNIFICANT HOMOLOGY TO THE ADENINE BINDING REGION OF
CC P-HYDROXYBENZOATE HYDROXYLASE FROM P. FLUORESCENS. IT SEEMS THAT
CC THIS PROTEIN BINDS ADENINE EITHER AS SUBSTRATE OR COFACTOR.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X02654; AAA92550.1; -.
CC PIR; A04497; Q0AG4T.
CC InterPro; IPR002937; Amino_Oxidase.
CC InterPro; IPR000624; Glycosidase.
CC InterPro; IPR000205; NAD_binding.
CC Pfam; PF01593; Amino_Oxidase; 1.
CC Pfam; PF02027; RoLB_RolC; 1.
CC Oxidoreductase; Monooxygenase; Auxin biosynthesis; Crown gall tumor;
KW T-DNA; Plasmid.
KW VARIANT 718 719 NR -> IQ (IN PTIA6NC).
FT VARIANT 721 721 P -> A (IN PTIA6NC).
FT SEQUENCE 755 AA; 83947 MW; 9FD2B83FEA001A4D CRC64;
SQ
Query Match 27.9%; Score 62; DB 1; Length 755;
Best Local Similarity 41.7%; Pred. No. 2.2;
Matches 15; Conservative 6; Mismatches 9; Indels 6; Gaps 1;

QY 7 DSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWIA 42
||| ::||| | | | | | | | | | | | | | |
DB 223 DSG-----RIGFFPDVPRPKVAVIGAGISGLVVA 252

RESULT 12
TR2N_AGRVI
ID TR2N_AGRVI STANDARD; PRT; 755 AA.
AC P25017;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tryptophan 2-monooxygenase (EC 1.13.12.3).
GN IAAM.
OS Agrobacterium vitis (Rhizobium vitis).
OG Plasmid pTITM4.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=373;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=TM4;
RX MEDLINE=91329707; PubMed=1868204;
RA Bonnard G., Vincent F., Otten L.;
RT "Sequence of Agrobacterium tumefaciens biotype III auxin genes.";
RL Plant Mol. Biol. 16:733-738(1991).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CG474;
RA Otten L., de Ruffray P.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: L-tryptophan + O(2) = indole-3-acetamide +
CC CO(2) + H(2)O.

```

```

CC
CC -!- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF AUXINS FROM TRYPTOPHAN.
CC
CC -!- SIMILARITY: SIGNIFICANT HOMOLOGY TO THE ADENINE BINDING REGION OF
CC P-HYDROXYBENZOATE HYDROXYLASE FROM P. FLUORESCENS. IT SEEMS THAT
CC THIS PROTEIN BINDS ADENINE EITHER AS SUBSTRATE OR COFACTOR.
CC
CC -!- SIMILARITY: STRONG, WITH ITS COUNTERPARTS FROM A.TUMEFACIENS
CC PLASMIDS PT115955, PTIACH5 AND PTIA6NC.
CC
CC -!- CAUTION: THE PLASMID PTITM4 CARRIES TWO T-REGIONS, THE TA AND
CC TB REGION, BOTH OF WHICH HAVE A FUNCTIONAL TAAM GENE, WITH LOW
CC HOMOLOGY BETWEEN THEM.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X56185; CAA39646.1; -.
CC EMBL; U83987; AAB41874.1; -.
CC PIR; S15450; DAAGMT.
CC PIR; S15002; S15002.
CC InterPro; IPR002937; Amino_Oxidase.
CC InterPro; IPR000624; Glycosidase.
CC InterPro; IPR000205; NAD_binding.
CC Pfam; PF01593; Amino_Oxidase; 1.
CC Pfam; PF02027; RoLB_RolC; 1.
CC Oxidoreductase; Monooxygenase; Auxin biosynthesis; Crown gall tumor;
KW T-DNA; Plasmid.
KW SEQUENCE 755 AA; 83972 MW; 6FA63E502343136F CRC64;
SQ
Query Match 27.9%; Score 62; DB 1; Length 755;
Best Local Similarity 40.5%; Pred. No. 2.2;
Matches 15; Conservative 7; Mismatches 9; Indels 6; Gaps 1;

QY 7 DSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWVIAT 43
||| ::||| | | | | | | | | | | | | | |
DB 223 DSG-----RIGFFPDVPRPKVAVIGAGISGLVVA 253

RESULT 13
POLG_PVYCH
ID POLG_PVYCH STANDARD; PRT; 327 AA.
AC P21294;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Nuclear inclusion protein B (NI-B) (NIB)
DE (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP)]
DE (Fragment).
OS Potato virus Y (strain Chinese isolate) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12218;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91016851; PubMed=2216735;
RA Zhou X.R., Fang R.X., Wang C.Q., Mang K.Q.;
RT "cDNA sequence of the 3'-coding region of PVY genome (the Chinese
RT isolate).";
RL Nucleic Acids Res. 18:5554-5554(1990).
CC
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC
CC -!- PPM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC
CC -!- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X54058; CAA37993.1; -;
 DR PIR: S11549; S11549.
 DR InterPro: IPR001592; Poty-coat.
 DR Pfam: PF00767; Poty-coat; 1.
 KW Transferase: RNA-directed RNA polymerase; Coat protein; Polyprotein.
 FT NON_TER 1
 FT CHAIN <1 60 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 61 327 COAT PROTEIN.
 SQ SEQUENCE 327 AA; 36868 MW; 8F8355E2DE6F2F18 CRC64;

Query Match 25.7%; Score 57; DB 1; Length 327;

Best Local Similarity 53.1%; Pred. No. 4.1;

Matches 17; Conservative 0; Mismatches 5; Indels 10; Gaps 3;

QY 1 DAEFRHDSGYEVHQLVFFAED---VGSNK 28

DB 47 DDEFEPDS-YEVHQQ-----ANDTIDAVGDNK 72

RESULT 14

Y226_MYCPN

ID Y226_MYCPN STANDARD; PRT; 503 AA.

AC P75462;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MG226 homolog (P10_orf503).

GN MPN319 OR MP517.

OS Mycoplasma pneumoniae.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2104;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 29342 / M129;

RX MEDLINE=97105885; PubMed=8948633;

RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,

RA Hermann R.;

RT "Complete sequence analysis of the genome of the bacterium Mycoplasma

RT pneumoniae."

RL Nucleic Acids Res. 24:4420-4449(1996).

CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -|- SIMILARITY: TO M.GENITALIUM MG225.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AE000051; AAB96165.1; -;
 CC InterPro: IPR002293; AA/rel.pmease1.
 CC InterPro: IPR004841; Permease.
 CC Pfam: PF00324; aa_permeases; 1.

CC Hypothetical protein; transmembrane; Complete proteome.
 CC TRANSMEM 20 40 POTENTIAL.
 CC TRANSMEM 43 63 POTENTIAL.
 CC TRANSMEM 106 126 POTENTIAL.
 CC TRANSMEM 138 158 POTENTIAL.
 CC TRANSMEM 166 186 POTENTIAL.
 CC TRANSMEM 215 235 POTENTIAL.
 CC TRANSMEM 249 269 POTENTIAL.
 CC TRANSMEM 301 321 POTENTIAL.
 CC TRANSMEM 359 379 POTENTIAL.
 CC TRANSMEM 405 425 POTENTIAL.
 CC TRANSMEM 443 463 POTENTIAL.
 CC TRANSMEM 468 488 POTENTIAL.
 CC SEQUENCE 503 AA; 54960 MW; 4BC1BFDE036985B2 CRC64;

Query Match 25.7%; Score 57; DB 1; Length 503;

Best Local Similarity 61.1%; Pred. No. 6.4;

Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 24 VGSNKGATIGLWGVGV 41

DB 131 VKDNGALIGLWGVGV 148

RESULT 15

SLGL_YEAST

ID SLGL_YEAST STANDARD; PRT; 378 AA.

AC P54867;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE SLG1 protein precursor.

GN SLG1 OR YOR008C OR UNF378.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA de Bettignies G., Bergez-Aulio P., Barthe C., Louvet O.,

RA Peypouquet M.F., Morel C., Dolignon F., Crouzet M.;

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

RP [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=97051599; PubMed=8896276;

RA Sterky F., Holmberg A., Pettersson B., Uhlen M.;

RT "The sequence of a 30 kb fragment on the left arm of chromosome XV

RT from Saccharomyces cerevisiae reveals 15 open reading frames, five of

RT which correspond to previously identified genes."

RL Yeast 12:1091-1095(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U39481; AAA85862.1; -;
 CC EMBL: U43491; AAC49488.1; -;
 CC EMBL: Z74916; CAA99196.1; -;
 CC SGD: S0005534; SLG1.
 CC InterPro: IPR002889; WSC.
 CC Pfam: PF01822; WSC; 1.
 CC SMART: SM00321; WSC; 1.
 CC Glycoprotein; Signal.
 CC SIGNAL 1 19 POTENTIAL.
 CC CHAIN 20 378 SLG1 PROTEIN.
 CC CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 378 AA; 39270 MW; EEE164F2374CCCE3 CRC64;

Query Match 25.5%; Score 56.5; DB 1; Length 378;

Best Local Similarity 42.4%; Pred. No. 5.5;

Matches 14; Conservative 5; Mismatches 5; Indels 9; Gaps 1;

QY 8 SGYEVHQLVFFAEDVGSNKGATIGLWGVGV 40

DB 251 SGSKTHKK-----ANVGAIVGVGVGV 274

Search completed: January 28, 2003, 14:14:55

Job time : 3.65774 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 2.1729 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-1
Perfect score: 222
Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNKGAIIGLMVGVVIAT 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	100.0	82	4	Q16014 homo sapien
2	222	100.0	82	4	Q16019 homo sapien
3	222	100.0	82	4	Q16020 homo sapien
4	222	100.0	82	4	P78438 homo sapien
5	222	100.0	534	13	Q93296 gallus gall
6	222	100.0	695	6	Q95KN7 macaca fasc
7	222	100.0	695	11	Q60496 cavia sp. p
8	222	100.0	695	13	Q9DGJ8 gallus gall
9	222	100.0	751	13	Q9DGJ7 gallus gall
10	222	100.0	770	6	Q9TU10 sus scrofa
11	217	97.7	569	13	Q9PVL1 gallus gall
12	212	95.5	97	4	Q13778 homo sapien
13	203	91.4	79	11	Q35463 cricetus
14	203	91.4	607	11	Q99K32 mus musculu
15	203	91.4	693	13	Q98SG0 xenopus lae
16	203	91.4	695	11	P97487 mus musculu

17	203	91.4	747	13	Q91963 xenopus. ep
18	200	90.1	695	13	Q98SF9 xenopus lae
19	193	86.9	699	13	O57394 narke japon
20	180	81.1	780	13	O73683 tetraodon f
21	176	79.3	33	4	Q9UC33 homo sapien
22	176	79.3	737	13	Q93279 fugu rubrip
23	162.5	73.2	357	13	Q8UUI8 brachydanio
24	162.5	73.2	472	13	Q8UUI8 brachydanio
25	162.5	73.2	612	13	Q919E7 brachydanio
26	162.5	73.2	738	13	Q90W28 brachydanio
27	162	73.0	30	4	Q9UCA9 homo sapien
28	161	72.5	239	13	Q8UUI7 brachydanio
29	161	72.5	694	13	Q8UUI9 brachydanio
30	147	66.2	28	4	Q9UCD1 homo sapien
31	126	56.8	49	6	Q97917 bos taurus
32	106	47.7	19	4	Q9UCB8 homo sapien
33	95	42.8	35	4	Q8WZ99 homo sapien
34	64	28.8	20	4	Q9UCB6 homo sapien
35	64	28.8	328	2	Q9RPS4 enterococcu
36	64	28.8	755	2	Q9R717 agrobacteri
37	64	28.8	755	2	Q9R472 agrobacteri
38	64	28.8	755	2	Q9R694 agrobacteri
39	64	28.8	755	16	Q8U6A3 agrobacteri
40	63	28.4	321	16	Q8RG41 fusobacteri
41	62	27.9	755	2	Q44388 agrobacteri
42	61	27.5	755	2	Q9WWA1 agrobacteri
43	57.5	25.9	895	10	Q9AWB6 lycopersico
44	57	25.7	195	10	O22662 arabidopsis
45	57	25.7	332	12	Q9DQN5 potato viru

ALIGNMENTS

RESULT 1

ID	Q16014	PRELIMINARY;	PRT;	82 AA.
AC	Q16014;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Beta-amyloid peptide (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93236601; PubMed=8476439;			
RA	Denman R.B., Rossenwaig R., Miller D.L.;			
RT	"A system for studying the effect(s) of familial Alzheimer disease mutations on the processing of the beta-amyloid peptide precursor."			
RL	Biochem. Biophys. Res. Commun. 192:96-103(1993).			
DR	EMBL; S60721; AAB26263.2; -.			
DR	HSSP; P05067; IBA4.			
DR	InterPro: IPR001255; Beta-APP.			
DR	Pfam: PF03494; Beta-APP; 1.			
FT	NON_TER	1		
FT	NON_TER	82		
SQ	SEQUENCE	82 AA;	8972 MW;	F534AA5B3EA9230A CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGVVIAT 43
|||||
Db 18 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGVVIAT 60
|||||

RESULT 2

ID	Q16019	PRELIMINARY;	PRT;	82 AA.
----	--------	--------------	------	--------


```

AC Q16019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid peptide (Fragment).
GN BETA APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzweig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL; S61380; AAB26264.2; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8938 MW; F534AA50E579230A CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 18 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 60

RESULT 3
Q16020
ID Q16020 PRELIMINARY; PRT; 82 AA.
AC Q16020;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid peptide (Fragment).
GN BETA APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzweig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
DR EMBL; S61383; AAB26265.2; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8882 MW; F534AA5AE5D9230A CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 18 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 60

RESULT 4
P78438

```

```

ID P78438 PRELIMINARY; PRT; 82 AA.
AC P78438;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (Beta-amyloid protein) (Fragment).
GN APP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=89392030; PubMed=2675837;
RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.;
RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN [2]
RP SEQUENCE OF 19-48 FROM N.A.
RP MEDLINE=87120329; PubMed=2949367;
RA Tanzi R.E., Gusella J.F., Watkins P.C., Bruns G.A., George-Hyslop P.,
RA Van Keuren M.L., Patterson D., Fagan S., Kurnit D.M., Neve R.L.;
RT "Amyloid beta protein gene: cDNA, mRNA distribution, and genetic
RT linkage near the Alzheimer locus.";
RL Science 235:880-884(1987).
RN [3]
RP SEQUENCE OF 32-63 FROM N.A.
RP MEDLINE=93035397; PubMed=1415269;
RA Kanino K., Orr H.T., Payami H., Wijsman E.M., Alonso M.E., Pulst S.M.,
RA Anderson L., O'dahl S., Nemens E., White J.A.;
RT "Linkage and mutational analysis of familial Alzheimer disease
RT kindreds for the App gene region.";
RL Am. J. Hum. Genet. 51:998-1014(1992).
DR EMBL; M29270; AAA51768.1; -.
DR EMBL; M29269; AAA51768.1; JOINED.
DR EMBL; M15532; AAA51564.1; -.
DR EMBL; S45136; AAB23646.1; -.
DR HSSP; P05067; 1BA4.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8994 MW; 8DA9E42B813A070E CRC64;

Query Match 100.0%; Score 222; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
Db 17 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 59

RESULT 5
Q93296
ID Q93296 PRELIMINARY; PRT; 534 AA.
AC Q93296;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid protein (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=98337885; PubMed=9671674;
RA Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
RA Milligan C.E.;
RT "Increased production of amyloid precursor protein provides a

```



```

RT substrate for caspase-3 in dying motoneurons.";
RL J. Neurosci. 18:5869-5880(1998).
DR EMBL: AF042098; AAC25052.1; -.
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PRINTS: PR00203; AMYLOIDA4.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
FT NON_TER 1
SQ SEQUENCE 534 AA; 60597 MW; FB53ECC2E66D4C92 CRC64;

Query Match 100.0%; Score 222; DB 13; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DQ 436 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 478

RESULT 6
Q95KN7 PRELIMINARY; PRT; 695 AA.
AC Q95KN7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amyloid b-protein precursor.
OS Macaca fascicularis (Crab eating macaque) (Cynomolpus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RX MEDLINE=91273117; PubMed=1905108;
RA Podlasky M.B., Tolan D.R., Selkoe D.J.;
RT "Homology of the amyloid beta protein precursor in monkey and human
RT supports a primate model for beta amyloidosis in Alzheimer's
RT disease.";
RL Am. J. Pathol. 138:1423-1435(1991).
DR EMBL: M58727; AAA36829.1; -.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; UNKNOWN_1.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 597 636 POTENTIAL.
SQ SEQUENCE 695 AA; 78663 MW; 4F6EA0139F969D56 CRC64;

Query Match 100.0%; Score 222; DB 6; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DQ 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

RESULT 7
Q60496 PRELIMINARY; PRT; 695 AA.
AC Q60496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative amyloid precursor protein.

```

```

OS Cavia sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10143;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97236426; PubMed=9116031;
RA Beck M., Mueller D., Bigl V.;
RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT alternative splicing.";
RL Biochim. Biophys. Acta 1351:17-21(1997).
DR EMBL: X97631; CAA66230.1; -.
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PRINTS: PR00203; AMYLOIDA4.
DR SMART: SM00006; A4_EXTRA; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

Query Match 100.0%; Score 222; DB 11; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DQ 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

RESULT 8
Q9DGJ8 PRELIMINARY; PRT; 695 AA.
AC Q9DGJ8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 695 Isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodollos A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
RT isoforms.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF289218; AAG00593.1; -.
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PRINTS: PR00203; AMYLOIDA4.
DR SMART: SM00006; A4_EXTRA; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match 100.0%; Score 222; DB 13; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.6e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DQ 597 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 639

```


PRINTS; PF00203; AMYLOIDA4.
 PRINTS; PR00759; BASICPTASE.
 PRODom; PD000222; Kunitz_BPTI; 1.
 SMART; SM00006; A4_EXTRA; 1.
 SMART; SM00131; KU; 1.
 PROSITE; PS00319; A4_EXTRA; 1.
 PROSITE; PS00320; A4_INTRA; 1.
 PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 PROSITE; PS0279; BPTI_KUNITZ_2; 1.
 Serine protease inhibitor.
 Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 Gallus.
 NCBI_TaxID=9031;
 [1]
 SEQUENCE FROM N.A.
 Sarasa M., Rodolosse A., Sorribas V.;
 "Cloning of full-length chicken beta-amyloid precursor protein
 isoforms";
 Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 EMBL; AF282219; RAG00594.1; -.
 HSSP; P05067; IBA4.
 InterPro; IPR001868; A4_APP.
 InterPro; IPR001255; Beta_APP.
 InterPro; IPR002223; Kunitz_BPTI.
 Pfam; PF02177; A4_EXTRA; 1.
 Pfam; PF03494; Beta_APP; 1.
 Pfam; PF00014; Kunitz_BPTI; 1.
 PRINTS; PR00203; AMYLOIDA4.
 PRINTS; PR00759; BASICPTASE.
 PRODom; PD000222; Kunitz_BPTI; 1.
 SMART; SM00006; A4_EXTRA; 1.
 SMART; SM00131; KU; 1.
 PROSITE; PS00319; A4_EXTRA; 1.
 PROSITE; PS00320; A4_INTRA; 1.
 PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 PROSITE; PS0279; BPTI_KUNITZ_2; 1.
 Serine protease inhibitor.
 SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;
 Query Match 100.0%; Score 222; DB 13; Length 751;
 Best Local Similarity 100.0%; Pred. No. 1.8e-21;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 653 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 695
 RESULT 10
 Q9TU10
 ID Q9TU10 PRELIMINARY; PRT; 770 AA.
 AC Q9TU10;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Amyloid precursor protein.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 [1]
 SEQUENCE FROM N.A.
 Kimura A., Takahashi T.;
 "Amyloid Precursor Protein 770.";
 Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 EMBL; AB032550; BAA84580.1; -.
 HSSP; P05067; 1AAP.
 InterPro; IPR001868; A4_APP.
 InterPro; IPR001255; Beta_APP.
 InterPro; IPR002223; Kunitz_BPTI.
 Pfam; PF02177; A4_EXTRA; 1.
 Pfam; PF03494; Beta_APP; 1.
 Pfam; PF00014; Kunitz_BPTI; 1.
 PRINTS; PR00203; AMYLOIDA4.
 PRINTS; PR00759; BASICPTASE.
 PRODom; PD000222; Kunitz_BPTI; 1.
 SMART; SM00006; A4_EXTRA; 1.
 SMART; SM00131; KU; 1.
 PROSITE; PS00319; A4_EXTRA; 1.
 PROSITE; PS00320; A4_INTRA; 1.
 PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 PROSITE; PS0279; BPTI_KUNITZ_2; 1.
 Serine protease inhibitor.
 SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;
 Query Match 100.0%; Score 222; DB 13; Length 751;
 Best Local Similarity 100.0%; Pred. No. 1.8e-21;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 653 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 695
 RESULT 9
 Q9DGJ7
 ID Q9DGJ7 PRELIMINARY; PRT; 751 AA.
 AC Q9DGJ7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Beta-amyloid precursor protein 751 isoform.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 Gallus.
 NCBI_TaxID=9031;
 [1]
 SEQUENCE FROM N.A.
 Sarasa M., Rodolosse A., Sorribas V.;
 "Cloning of full-length chicken beta-amyloid precursor protein
 isoforms";
 Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 EMBL; AF282219; RAG00594.1; -.
 HSSP; P05067; IBA4.
 InterPro; IPR001868; A4_APP.
 InterPro; IPR001255; Beta_APP.
 InterPro; IPR002223; Kunitz_BPTI.
 Pfam; PF02177; A4_EXTRA; 1.
 Pfam; PF03494; Beta_APP; 1.
 Pfam; PF00014; Kunitz_BPTI; 1.
 PRINTS; PR00203; AMYLOIDA4.
 PRINTS; PR00759; BASICPTASE.
 PRODom; PD000222; Kunitz_BPTI; 1.
 SMART; SM00006; A4_EXTRA; 1.
 SMART; SM00131; KU; 1.
 PROSITE; PS00319; A4_EXTRA; 1.
 PROSITE; PS00320; A4_INTRA; 1.
 PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 PROSITE; PS0279; BPTI_KUNITZ_2; 1.
 Serine protease inhibitor.
 SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;
 Query Match 100.0%; Score 222; DB 6; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1.8e-21;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 DB 672 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 714
 RESULT 11
 Q9PVL1
 ID Q9PVL1 PRELIMINARY; PRT; 569 AA.
 AC Q9PVL1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Amyloid protein (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 Gallus.
 NCBI_TaxID=9031;
 [1]
 SEQUENCE FROM N.A.
 Coulson E.J., Paliga K., Beyreuther K., Masters C.L.;
 "What the evolution of the amyloid protein precursor superegene family
 tells us about its function.";
 Neurochem Int. 0:0-0(2000).
 EMBL; AF030341; AAF12698.1; -.
 HSSP; P05067; IBA4.
 InterPro; IPR001868; A4_APP.
 InterPro; IPR001255; Beta_APP.
 Pfam; PF02177; A4_EXTRA; 1.
 Pfam; PF03494; Beta_APP; 1.
 PRINTS; PR00203; AMYLOIDA4.
 SMART; SM00006; A4_EXTRA; 1.
 PROSITE; PS00319; A4_EXTRA; 1.
 PROSITE; PS00320; A4_INTRA; 1.
 NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 569 AA; 64753 MW; 0AB8BB851863A19D CRC64;
 Query Match 97.7%; Score 217; DB 13; Length 569;
 Best Local Similarity 100.0%; Pred. No. 6e-21;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 42
 DB 472 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 513
 RESULT 12
 Q13778
 ID Q13778 PRELIMINARY; PRT; 97 AA.
 AC Q13778;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Amyloid protein (AD-AP) (Fragment).
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87120328; PubMed=3810169;
 RA Goldhaber D., Lerman M.I., McBride O.W., Saffioti U., Gajdusek D.C.;
 RT "Characterization and chromosomal localization of a cDNA encoding
 brain amyloid of Alzheimer's disease.";
 RL Science 235:877-880(1987).
 DR EMBL: M15533; AAA35540.1; -.
 DR HSSP: P05067; IBA4.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR001255; Beta-APP.
 DR Pfam: PF03494; Beta-APP; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 FT NON_TER 1
 SQ SEQUENCE 97 AA; 10884 MW; E528CDB448DE474E CRC64;
 Query Match 95.5%; Score 212; DB 4; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.5e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 EFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 1 EFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 41
 RESULT 13
 O35463
 ID O35463 PRELIMINARY; PRT; 79 AA.
 AC O35463;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Alzheimer's amyloid beta protein (Fragment).
 GN BETA APP.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sambamurti K., Pinnix I., Gandhi S.;
 RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF030413; AAB86608.1; -.
 DR HSSP: P05067; IBA4.
 DR InterPro: IPR001255; Beta-APP.
 DR Pfam: PF03494; Beta-APP; 1.
 FT NON_TER 1
 FT NON_TER 79
 SQ SEQUENCE 79 AA; 8538 MW; 37F2C6C3BF3F597 CRC64;
 Query Match 91.4%; Score 203; DB 11; Length 79;
 Best Local Similarity 93.0%; Pred. No. 4.6e-20;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 21 DAEFGHDSGFVRRHQLVFFAEDVGSNKGAIIGLMVGGVVAT 63
 RESULT 14
 Q99K32
 ID Q99K32 PRELIMINARY; PRT; 607 AA.
 AC Q99K32;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 68.4 kDa protein (Fragment).
 GN APP.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC005490; AAH05490.1; -.
 DR HSSP: P05067; IAAAP.
 DR MGD: MGI:88059; App.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR001255; Beta-APP.
 DR InterPro: IPR002223; Kunitz_BPTI.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR Pfam: PF03494; Beta-APP; 1.
 DR Pfam: PF00014; Kunitz_BPTI; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR PRINTS: PR00759; BASICPTASE.
 DR PRODOM: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00131; KU; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS02079; BPTI_KUNITZ_2; 1.
 KW Hypothetical protein; Serine protease inhibitor.
 FT NON_TER 1
 SQ SEQUENCE 607 AA; 68391 MW; BF802214CBA7D172 CRC64;
 Query Match 91.4%; Score 203; DB 11; Length 607;
 Best Local Similarity 93.0%; Pred. No. 5.2e-19;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 Db 509 DAEFGHDSGFVRRHQLVFFAEDVGSNKGAIIGLMVGGVVAT 551
 RESULT 15
 Q98SG0
 ID Q98SG0 PRELIMINARY; PRT; 693 AA.
 AC Q98SG0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Beta-amyloid precursor protein A.
 GN APP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA van den Hurk W.H.;
 RL Thesis (2001), Department of Biological Sciences,
 University of Nijmegen, Nijmegen, Netherlands.
 DR EMBL: AJ298150; CAC37193.1; -.
 DR HSSP: P05067; IHZ3.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR001255; Beta-APP.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR Pfam: PF03494; Beta-APP; 1.
 DR PRINTS: PR00203; AMYLOIDA4.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 693 AA; 78568 MW; CAF1DF655C1AB653 CRC64;
 Query Match 91.4%; Score 203; DB 13; Length 693;
 Best Local Similarity 88.4%; Pred. No. 6.1e-19;
 Matches 38; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Search completed: January 28, 2003, 14:18:11
Job time : 8.1729 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 : Search time 2.39607 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEPRHDSGEVHHQKLVF.....VGSNGKALICLMVGGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222	100.0	43	15	AA54759
2	222	100.0	43	15	AA60367
3	222	100.0	43	15	AA61328
4	222	100.0	43	16	AA64165
5	222	100.0	43	17	AA69573
6	222	100.0	43	18	AA93371
7	222	100.0	43	19	AA17758
8	222	100.0	43	19	AA71378
9	222	100.0	43	19	AA51316
10	222	100.0	43	19	AA40129

11	222	100.0	43	20	AA42955	Beta-amyloid precu
12	222	100.0	43	20	AAW89362	Beta-amyloid pepti
13	222	100.0	43	21	AA527020	Beta-amyloid pepti
14	222	100.0	43	21	AA515372	Human beta-amyloid
15	222	100.0	43	21	AA521216	Beta-amyloid pepti
16	222	100.0	43	21	AA58390	Beta-amyloid pepti
17	222	100.0	43	21	AA556102	Natural beta amylo
18	222	100.0	43	22	AB07901	Beta-amyloid pepti
19	222	100.0	43	22	AA678791	Human beta amyloid
20	222	100.0	43	22	AAE12508	Beta-amyloid pepti
21	222	100.0	43	22	AA84428	Partial sequence o
22	222	100.0	43	22	AA89886	Beta-amyloid pepti
23	222	100.0	43	22	AA81193	Beta-amyloid pepti
24	222	100.0	43	22	AA891778	Amyloid beta-prote
25	222	100.0	43	22	AA891811	Amyloid beta-prote
26	222	100.0	43	22	AA847108	Biotinylated beta-
27	222	100.0	43	22	AA848344	Beta-amyloid pepti
28	222	100.0	43	23	AA98701	Human amyloid beta
29	222	100.0	43	23	AA50862	Beta-amyloid pepti
30	222	100.0	43	23	AB05149	Beta amyloid pepti
31	222	100.0	45	16	AA64169	Variant beta amylo
32	222	100.0	47	20	AA81475	Synthetic amyloid
33	222	100.0	48	22	AA537523	Amyloid precursor
34	222	100.0	52	16	AA64166	Variant beta amylo
35	222	100.0	52	20	AA81476	Synthetic amyloid
36	222	100.0	53	15	AA55695	Sequence of uniden
37	222	100.0	53	16	AA64168	Variant beta amylo
38	222	100.0	54	21	AA32126	Amyloid-beta precu
39	222	100.0	55	22	AA811482	Human APP peptide
40	222	100.0	57	21	AA810910	Human amyloid prec
41	222	100.0	59	17	AAW05375	Amyloid precursor
42	222	100.0	59	19	AAW70863	Beta-amyloid precu
43	222	100.0	59	22	AA84425	Partial sequence o
44	222	100.0	60	21	AA69701	Beta-amyloid precu
45	222	100.0	63	19	AA44747	APP-REP 751 BAP pe

ALIGNMENTS

RESULT 1

AA54759

ID AA54759 standard; peptide; 43 AA.

XX AA54759;

AC AA54759;

DT 30-NOV-1994 (first entry)

XX Beta amyloid peptide.

DE Beta amyloid peptide.

XX Beta amyloid peptide; amyloid plaques; Alzheimer's disease; lesion; brain; senility; dementia; detection; diagnosis.

XX Homo sapiens.

XX WO9410569-A.

XX 11-MAY-1994.

XX 01-SEP-1993; 93WO-US08264.

XX 26-OCT-1992; 92US-0965972.

XX (SCHE/) SCHENK D B.

XX (SCHL/) SCHLOSSMACHER M G.

XX (SELK/) SELKOE D.

XX Schenk DB, Schlossmacher MG, Selkoe DJ, Seubert PA;

XX Vigo-pelfrey;

XX WPI; 1994-167654/20.

XX Detecting soluble beta-amyloid peptide concns. e.g. for

PT diagnosing and assessing progression of Alzheimer's disease - by
 PT exposing cultured cells to test cpd. to determine effect of cpd.
 PT on produced soluble beta-amyloid peptide

XX Disclosure; Page 10; 55pp; English.

XX
 XX Beta amyloid peptide is the principal chemical constituent of
 CC amyloid plaques, lesions found on the brains of Alzheimer's disease
 CC patients. The ability to detect beta amyloid peptide in fluid
 CC samples provides a means of diagnosing Alzheimer's disease.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 2

AAR60367.
 ID AAR60367 standard; peptide; 43 AA.

XX AAR60367;

XX 15-MAR-1995 (first entry)

XX Beta-amyloid (1-43).

XX Amyloid precursor protein; APP; Alzheimer's disease; beta-amyloid;
 KW anti-beta-amyloid antibody; diagnosis.

XX Homo sapiens.

XX WO9417197-A.

XX 04-AUG-1994.

XX 24-JAN-1994; 94WO-JP00089.

XX 25-JAN-1993; 93JP-0010132.

XX 05-FEB-1993; 93JP-0019035.

XX 16-NOV-1993; 93JP-0286985.

XX 28-DEC-1993; 93JP-0334773.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Kitada C, Odaka A, Suzuki N;

XX WPI; 1994-264110/32.

XX Antibodies recognising specific parts of beta-amyloid - can be
 PT used for diagnosis of diseases implicating beta-amyloid, such as
 PT Alzheimer's disease

XX Disclosure; Page 83; 116pp; Japanese.

XX Antibodies which recognise specific subfragments of the beta-amyloid
 CC protein are claimed. Specifically, the antibodies (which are pref.
 CC monoclonal) recognise residues 1-16 and/or 1-28 from the N-terminal
 CC portion of beta-amyloid or they recognise residues 25-35 or 35-43
 CC from the C-terminal portion. The antibodies are useful for assaying
 CC beta-amyloid and its derivatives for diagnosis of Alzheimer's
 CC disease.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 3

AAR61328
 ID AAR61328 standard; Protein; 43 AA.

XX AAR61328;

XX 21-APR-1995 (first entry)

XX Amyloid beta-protein, ABP.

XX Amyloid beta-protein; ABP; Tau-protein kinase I enzyme; TPK-I;

KW Alzheimer's disease.

XX Synthetic.

XX EP616032-A.

XX 21-SEP-1994.

XX 01-MAR-1994; 94EP-0103057.

XX 02-MAR-1993; 93JP-0041160.

XX 22-MAR-1993; 93JP-0085143.

XX 02-AUG-1993; 93JP-0191246.

XX (TAKA/) TAKASHIMA A.

PA (MITU) MITSUBISHI KASEI CORP.

XX Hoshino T, Imahori K, Saito K, Sato S, Shiratsuchi A;

PI Takashima A;

XX WPI; 1994-287181/36.

XX Newly isolated tau-protein kinase I enzyme - with specificity for
 PT tau-protein providing means for prevention and treatment of
 PT Alzheimer's disease

XX Example 1; Page 22; 30pp; English.

XX Amyloid beta-protein (ABP) is the main component of senile plaques
 CC in Alzheimer's disease. ABP was used to demonstrate the protective
 CC properties of anti-sense TPK-I oligonucleotides (AAQ67461 and AAQ67462)
 CC on hippocampus cells dosed with ABP.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 15; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 4

AAR64165
 ID AAR64165 standard; peptide; 43 AA.

XX AAR64165;

XX 02-AUG-1995 (first entry)

XX Beta amyloid protein.

XX beta amyloid protein; mutant; variant; detection; amyloid deposition;
 KW diagnosis; amyloidosis associated disease; Alzheimer's disease;

KW Down's syndrome.
 OS Synthetic.
 XX WO9428412-A.
 PN 08-DEC-1994.
 XX 27-MAY-1994; 94WO-US05809.
 PF 28-MAY-1993; 93US-0069010.
 PR (MIRI-) MIRIAM HOSPITAL.
 PA Majocha RE, Marotta CA;
 PI WPI; 1995-023013/03.
 DR
 XX Amyloid binding composition comprising labelled amyloid protein
 PT and carrier - useful for in vivo imaging of amyloid deposits, for
 PT diagnosing Alzheimer's disease and Down's Syndrome.
 PS Claim 5; Page 42; 58pp; English.
 XX AAR64165 shows the amino acid sequence of the beta amyloid protein. The
 CC protein binds amyloid and is useful for in vivo imaging of amyloid
 CC deposits and hence diagnosis of an amyloidosis-associated disease, such
 CC as Alzheimer's disease or Down's syndrome. AAR64165-69 show specific
 CC variants generated from this generic sequence with addition amino acids.
 XX
 SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 16; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 5
 AAR95673 standard; Peptide; 43 AA.
 XX
 AC AAR95673;
 XX
 DT 24-FEB-1997 (first entry)
 XX
 DE A-beta protein (43 amino acid version).
 XX
 KW A-beta protein; beta-amyloid precursor protein; APP; Alzheimer's disease;
 KW diagnosis; monitor; amyloid plaque; senile.
 XX
 OS Homo sapiens.
 XX
 PN WO9615452-A1.
 XX
 PD 23-MAY-1996.
 XX
 PF 13-NOV-1995; 95WO-US14659.
 XX
 PR 07-APR-1995; 95US-0419008.
 PR 14-NOV-1994; 94US-0339141.
 XX
 XX (ATHE-) ATHENA NEUROSCIENCES INC.
 XX
 PI Barbour R, Schenk DB, Seubert PA, Vigo-pelfrey C;
 XX
 DR WPI; 1996-260003/26.
 XX
 XX Diagnosis and monitoring of Alzheimer's disease - by detecting
 PT abnormally low concentration of A-beta peptide extending beyond

PT amino acid 41 in cerebrospinal fluid
 XX
 PS Disclosure; Page 16; 57pp; English.
 XX
 CC The "amyloid-beta peptide" or A-beta is an approx. 4.2 kD protein which,
 CC in the brains of Alzheimer's disease, Down's Syndrome, HCHWA-D and some
 CC normal aged subjects, forms the subunit of amyloid filaments comprising
 CC the senile (amyloid) plaques and the amyloid deposits in small cerebral
 CC and meningeal blood vessels. A-beta is an approx. 39-43 amino acid
 CC fragment of a large membrane-spanning glycoprotein, referred to as the
 CC beta-amyloid precursor protein (APP), encoded by a gene on the long arm
 CC of human chromosome 21. Detecting the amt. of A-beta is useful in
 CC diagnosis and monitoring of Alzheimer's disease, when taken together
 CC with other clinical symptoms. The present sequence is a 43 amino acid
 CC version of A-beta.
 XX
 SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 17; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAI 43
 RESULT 6
 AAW93371
 ID AAW93371 standard; Protein; 43 AA.
 XX
 AC AAW93371;
 XX
 DT 28-MAY-1999 (first entry)
 XX
 DE Human beta-amyloid polypeptide.
 XX
 KW Cathepsin Y; human; beta-amyloid peptide; BAP; secretion; inhibitor;
 KW protease; carboxypeptidase; aliphatic C-terminal amino acid; diagnostic;
 KW glycoprotein; cerebral deposition; pathogenesis; Alzheimer's disease;
 KW detection; amyloid plaque; angiopathy; brain; Trisomy 21; amyloidosis;
 KW Down's syndrome; hereditary cerebral haemorrhage.
 XX
 OS Homo sapiens.
 XX
 PN WO9639194-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 26-APR-1996; 96WO-US06211.
 XX
 PR 06-JUN-1995; 95US-0469362.
 PR 06-JUN-1995; 95US-0467607.
 PR 02-MAY-1997; 97US-0850392.
 XX
 XX (ATHE-) ATHENA NEUROSCIENCES INC.
 PA (ANDE-) ANDERSON J.
 PA (CHRY-) CHRISLER S.
 PA (MCCO-) MCCONLOGUE L.
 PA (SINH-) SINHA S.
 PA (TATS-) TATSUNO G.
 PA (TUNG-) TUNG JS.
 XX
 PI Anderson J, Chrysler S, McConlogue L, Semko CMF;
 PI Sinha S, Tatsuno G, Tung JS;
 XX
 DR WPI; 1997-042872/04.
 XX
 PT Acylamino and acyl-peptido-amino alcohol and aldehyde derivs. -
 PT inhibit beta-amyloid peptide prodn. in cells, use in Alzheimer's
 PT disease, also prepn. of cathepsin Y and nucleic acid encoding for
 PT it.
 XX

PS Disclosure; Page 11; 90pp; English.

XX This invention describes the inhibition of beta-amyloid peptide
 CC production in cells is effected by administration of an acylamino or
 CC acylpeptidoamino alcohol or aldehyde derivative of formula
 CC R1(X)m-Y-NR2-CH(R3)-R4 where R and R10 = H or 1-6C alkyl;
 CC or R and R2 together, and/or R10 and R3 together complete a 4-10C ring
 CC structure; R1 = (a) 1-4C alkyl substd. by 1-5 substituents chosen from 6-10C
 CC aryl (opt. substd. by 1-3 of 1-6C alkyl, 1-6C alkoxy, 6-10C aryl, 6-10C
 CC aryloxy, OH, cyano, halo and amino), 3-8C cycloalkyl or Het, in which the
 CC substd. alkyl gp. is opt. further substd. by 1 or 2 of OH; (b) 2-4C
 CC alkenyl substd. by 1-4 substituents, as for 1-4C alkyl above; (c) 6-10C aryl
 CC (opt. substd. by 1-3 substituents, as for alkyl above), (d) fluorenyl or (e)
 CC Het; Het = 3-14C heterocyclyl contg. 1-3 N, O or S heteroatoms; R2 and R3
 CC = D- or L- amino acid side chains of at least 2C atoms, excluding prolyl
 CC side chain; R4 = COCH=N2, CH2OH, C=NOH or COR5; R5 = H, 1-6C alkyl (opt.
 CC contg. 1 or 2 halogen atoms), 1-6C alkoxy, NR6R7 or N(Me)OMe; R6 and R7
 CC = H or 1-6C alkyl; X = O, NR9 or S; R9 = H, 1-6C alkyl or 6-10C aryl; Y =
 CC CO or CS; m = 0 or 1; and n = 0-2; provided that (i) when R1 =
 CC 1-naphthyl, R2 = CHMe2 (L-isomer), R3 = benzyl (L-isomer), Y = CO, m =
 CC 0 and n = 1, then R4 (sic) is not N(Me)OMe; (ii) when R1 = CHPh2, R2 =
 CC p-benzoyloxybenzyl (L-isomer), Y = CO, and m = n = 0, then R4 (sic) is
 CC not N(Me)OMe; and (iii) when R1 = PhCH=CPh, Y = CO, R2 = benzyl
 CC (L-isomer) and m = n = 0, then R4 (sic) is not N(Me)OMe. Cathepsin Y is
 CC a 31 kD carboxypeptidase with particular activity for aliphatic
 CC C-terminal amino acids. It is involved in the generation of beta-amyloid
 CC protein (BAP) from its precursor glycoprotein. Cerebral deposition of BAP
 CC plays a key role in pathogenesis of Alzheimer's disease (AD), and its
 CC detection can precede openly apparent symptoms of AD by years or more.
 CC Amyloid plaques and angiopathy are also present in the brain in those
 CC having Trisomy 21 (Down's syndrome) and hereditary cerebral haemorrhage
 CC with amyloidosis of Dutch type (HCHWA-D). The inhibitors are of use in
 CC inhibiting such plaque deposition and in treating AD. The cathepsin Y can
 CC be used, by virtue of its carboxypeptidase activity, in screening for BAP
 CC prodn. inhibitors, and, more widely, as a general C-terminal protease in
 CC a variety of chemical and biological systems. Conversely, assay of
 CC cathepsin Y in body fluid can be used diagnostically to evaluate the
 CC future risk of developing AD. The nucleic acid can be used to probe
 CC specifically for presence of cathepsin Y RNA or DNA in tissues or cloned
 CC libraries. It can also be used for expression of recombinant cathepsin Y,
 CC by transfection of a host cell.

XX
 SQ Sequence 43 AA;

Query Match 100.0%; Score 222; DB 18; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||

RESULT 7
 AAY17758
 ID AAY17758 standard; peptide: 43 AA.
 XX
 AC AAY17758;
 DT 10-AUG-1999 (first entry)
 XX
 DE Beta-amyloid peptide.
 XX
 KW Beta-amyloid peptide; beta-amyloid precursor protein; APP;
 KW membrane spanning glycoprotein; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9822493-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 20-NOV-1997; 97WO-US18704.

XX
 PR 22-NOV-1996; 96US-0755334.
 XX
 PA (ATHE-) ATHENA NEUROSCIENCES INC.
 PA (ELLIL) LILLY & CO ELI.
 XX
 PI Audia JE, Folmer BK, John V, Latimer LH, Nissen JS;
 PI Porter WJ, Thorsett ED, Wu J;
 XX
 WPI; 1998-322359/28.
 DR
 XX New and known aryl or hetero-aryl amino acid derivatives - useful
 PT for inhibiting beta-amyloid peptide release and/or its synthesis and
 PT treating Alzheimer's disease
 XX
 PS Disclosure; Page 19; 131pp; English.
 XX
 CC The present invention describes a composition comprising an inert carrier
 CC and an N-(aryl or heteroaryl) amino acid derivative e.g.
 CC N-(N-(3,4-dichlorophenyl)alanyl) valine methyl ester. The composition
 CC can be used for inhibiting beta-amyloid peptide release and/or its
 CC synthesis in a cell, for preventing the onset of Alzheimer's disease
 CC and for treating Alzheimer's disease in order to inhibit further
 CC deterioration. The dosage is 0.1-500 mg/kg/day orally, rectally,
 CC transdermally, subcutaneously or intravenously. The present sequence
 CC represents the beta-amyloid peptide.
 XX
 SQ Sequence 43 AA;

Query Match 100.0%; Score 222; DB 19; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||

RESULT 8
 AAW71378
 ID AAW71378 standard; peptide: 43 AA.
 XX
 AC AAW71378;
 DT 02-DEC-1998 (first entry)
 XX
 DE Beta-amyloid precursor protein epitope.
 XX
 KW Beta amyloid precursor protein; eukaryotic cell line;
 KW exogenous gene construction; identification; quantification;
 KW inhibitor; beta-amyloid processing; treatment; Alzheimer's disease;
 KW brain trauma; Downs syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO9837215-A1.
 XX
 PD 27-AUG-1998.
 XX
 PF 03-FEB-1998; 98WO-US01899.
 XX
 PR 31-JUL-1997; 97US-0904296.
 PR 24-FEB-1997; 97US-0804971.
 PR 02-APR-1997; 97US-0825737.
 XX
 PA (HMRI) HOECHST MARION ROUSSEL INC.
 PA (SCIO-) SCIOS INC.
 XX
 PI Cordell B, Huggins J, Mischak RP, Pruss R, Rautmann G;
 PI Scordina JM;
 XX
 DR WPI; 1998-495368/42.

CC deposition such as Down's syndrome, hereditary cerebral haemorrhage with
 CC amyloidosis of Dutch type and advanced aging of the brain.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 19; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 11

AA42955
 ID AAY42955 standard; protein; 43 AA.

XX
 AC AAY42955;

XX 02-JAN-2000 (first entry)

DE Beta-amyloid precursor protein.

XX Beta-amyloid precursor protein protein; APP; release inhibition;
 KW Alzheimer's disease.

XX Homo sapiens.

PN US5965614-A.

XX 12-OCT-1999.

XX 21-NOV-1997; 97US-0975977.

XX 22-NOV-1996; 96US-0104593.

XX (ATHE-) ATHENA NEUROSCIENCES INC.

XX Folmer BK, Audia JE, Latimer LH, Nissen JS, Reel JK, Thorsett ED;
 PI Whitesitt CA, John V;

DR WPI; 1999-579933/49.

XX Process for inhibiting beta-amyloid peptide cellular release and/or
 PT synthesis comprising administration of N-(aryl/heteroaryl) amino acid
 PT ester derivatives, useful for treatment and prevention of Alzheimer's
 PT disease -

XX Disclosure; Column 49; 29pp; English.

XX New chemical compounds are disclosed which inhibit the cellular
 CC release and/or synthesis of beta-amyloid peptide (the present sequence).
 CC The compounds are N-(aryl or heteroaryl) amino acid ester derivatives, of
 CC which N-(3,4-dichlorophenyl)alanine ethyl ester is a specific example.
 CC These compounds can be used for the treatment and prevention of
 CC Alzheimer's disease.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 20; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 12

AAW89362

ID AAW89362 standard; peptide; 43 AA.

XX

AC AAW89362;

XX 02-MAR-1999 (first entry)

DE Beta-amyloid peptide derivative A-beta-1-43.

XX Human; beta-amyloid peptide; Alzheimer's disease; amyloidogenic protein;
 KW aggregation; neurotoxicity; amyloidosis; Down's syndrome; cardiomyopathy;
 KW familial amyloid polynuropathy; bovine spongiform encephalopathy;
 KW Creutzfeldt-Jakob disease; BAP.

XX Homo sapiens.

OS Synthetic.

XX US5854204-A.

XX 29-DEC-1998.

XX 14-MAR-1996; 96US-0612785.

XX 14-MAR-1996; 96US-0612785.

XX 14-MAR-1995; 95US-0404831.

XX 07-JUN-1995; 95US-0475579.

XX 27-OCT-1995; 95US-0548998.

XX (PRAE-) PRAECIS PHARM INC.

XX Benjamin H, Chin J, Findels MA, Garnick MB, Geffer ML;
 PI Hundal A, Kasman L, Kelley M, Kubasek W, Lee J;

PI Molineaux S, Musso G, Reed M, Signer ER, Wakefield J;

XX WPI; 1999-094964/08.

XX New peptide(s) derived from beta-amyloid peptide that inhibit
 PT amyloid aggregation - and neurotoxicity, specifically for treatment
 PT and prevention of Alzheimer's disease

XX Example 1; Column 46; 52pp; English.

XX The present invention describes beta-amyloid peptide (BAP) derivatives.
 CC The BAP derivatives inhibit aggregation of amyloidogenic proteins and
 CC peptides, specifically BAP, and their neurotoxicity, so are useful for
 CC treating and preventing any disease involving amyloidosis, specifically
 CC Alzheimer's disease but also Down's syndrome, familial amyloid
 CC polynuropathy or cardiomyopathy, bovine spongiform encephalopathy and
 CC Creutzfeldt-Jakob disease. The BAP derivatives are also used to diagnose
 CC these diseases, in vitro or in vivo, by detecting binding of BAP to
 CC labelled BAP derivatives. Some BAP derivatives inhibit BAP aggregation
 CC even when BAP is present in molar excess. The present sequence
 CC represents a BAP derivative.

XX Sequence 43 AA;

Query Match 100.0%; Score 222; DB 20; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
 |||||
 Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 13

AA27020

ID AAB27020 standard; protein; 43 AA.

XX AAB27020;

XX 02-FEB-2001 (first entry)

XX Beta-amyloid peptide.

XX Beta-amyloid peptide; neurotropic; neuroprotective; Alzheimer's disease.

XX OS Unidentified.
 XX PN WO200052048-A1.
 XX PD 08-SEP-2000.
 XX PF 03-MAR-2000; 2000WO-US05574.
 XX PR 04-MAR-1999; 99US-0122736.
 XX PA (PRAE-) PRAECIS PHARM INC.
 XX PI Findeis MA, Phillips K, Olson GL, Self C;
 XX DR WPI; 2000-594168/56.
 XX PT Novel compounds that are useful as modulators of beta-amyloid peptide
 XX aggregation in treating amyloidosis, comprises D-amino acids -
 XX PS Disclosure; Page 9; 87pp; English.
 XX CC The present sequence is beta-amyloid peptide.. The present invention
 CC relates to peptides (see AAB27023-B27046) that modulate beta-amyloid
 CC peptide aggregation, and hence inhibit the neurotoxicity of beta-amyloid
 CC peptide. The beta-amyloid peptide modulators would be useful for
 CC treating disorders associated with beta-amyloidosis for e.g. Alzheimer's
 CC disease. The present sequence is derived from amyloid precursor protein
 CC via proteolysis. The gene for amyloid precursor protein maps to
 CC chromosome 21.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 43
 DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 43
 RESULT 14
 AAB15372
 ID AAB15372 standard; peptide; 43 AA.
 AC AAB15372;
 DT 17-JAN-2001 (first entry)
 DE Human beta-amyloid precursor protein A-beta fragment.
 XX Human; beta-amyloid precursor protein; beta-APP; Alzheimer's disease;
 XX chromosome 21; epitope.
 XX OS Homo sapiens.
 XX PN WO200042166-A2.
 XX PD 20-JUL-2000.
 XX PF 13-JAN-2000; 2000WO-US00872.
 XX PR 13-JAN-1999; 99US-0115749.
 XX PA (DUPO) DUPONT PHARM CO.
 XX PI Seiffert DA, Mitchell TJ;
 XX DR WPI; 2000-476049/41.
 XX PT A CDNA construct that encodes beta-amyloid precursor protein for
 XX identifying compounds which inhibit A-beta peptide release and/or

PT synthesis comprises an epitope tag within the A-beta sequence -
 XX Disclosure; Page 13-14; 42pp; English.
 XX CC The present sequence is A-beta fragment of the human beta-amyloid
 CC precursor protein (beta-APP). It was used in the production of peptides
 CC of the invention. The invention concerns the production, detection and
 CC characterisation of epitope-tagged beta-APP proteins, and their use in
 CC identifying modulators of beta-APP which can be used to treat diseases
 CC associated with an altered metabolism of the protein, in particular
 CC Alzheimer's disease.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 43
 DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA 43
 RESULT 15
 AAB21216
 ID AAB21216 standard; peptide; 43 AA.
 AC AAB21216;
 DT 11-JAN-2001 (first entry)
 DE Beta-amyloid peptide.
 XX Beta-amyloid peptide; membrane-spanning glycoprotein; analgesic;
 KW nicotinic agonist; beta-amyloid precursor protein; APP;
 KW Alzheimer's disease.
 XX OS Homo sapiens.
 XX PN US6117901-A.
 XX PD 12-SEP-2000.
 XX PF 21-NOV-1997; 97US-0976179.
 XX PR 22-NOV-1996; 96US-0098551.
 XX PA (ATHE-) ATHENA NEUROSCIENCES INC.
 XX PA (ELIL) LILLY & CO ELI.
 XX PI Thorsett ED, Nissen JS, Wu J, Latimer LH, John V, Fang LY;
 XX PI Audia JE, Mabry TE;
 XX DR WPI; 2000-637551/61.
 XX PT Eliciting analgesic effect in mammal, e.g. human, involves
 XX administering an aryl substituted olefinic amine compound -
 XX PS Disclosure; Column 11; 32pp; English.
 XX CC The present sequence is the beta-amyloid peptide, which is part of
 CC a large membrane-spanning glycoprotein, referred to as the beta-amyloid
 CC precursor protein (APP). Aryl substituted olefinic amine (metanicoline)
 CC compounds which inhibit beta-amyloid peptide release and/or its
 CC synthesis may be useful for treating Alzheimer's disease, both
 CC prophylactically and therapeutically.
 XX SQ Sequence 43 AA;
 Query Match 100.0%; Score 222; DB 21; Length 43;
 Best Local Similarity 100.0%; Pred. No. 1.9e-25;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

Search completed: January 28, 2003, 14:13:52
Job time : 6.39607 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 : Search time 0.56378 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGYEVHHQKLVFF.....VGSNKGAIIGLMVGGWIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	222	100.0	43	US-10-076-708-7	Sequence 7, Appli
2	222	100.0	43	US-10-051-496-1	Sequence 1, Appli
3	222	100.0	43	US-09-280-966-1	Sequence 1, Appli
4	222	100.0	43	US-09-904-987-1	Sequence 1, Appli
5	222	100.0	43	US-09-808-037-3	Sequence 3, Appli
6	222	100.0	43	US-09-866-712-3	Sequence 3, Appli
7	222	100.0	43	US-09-972-475-1	Sequence 3, Appli
8	222	100.0	43	US-09-982-800-1	Sequence 1, Appli
9	222	100.0	43	US-09-895-443-1	Sequence 1, Appli
10	222	100.0	43	US-09-996-357-1	Sequence 1, Appli
11	222	100.0	43	US-09-982-994-1	Sequence 1, Appli
12	222	100.0	43	US-09-984-834-1	Sequence 1, Appli
13	222	100.0	43	US-10-041-605-1	Sequence 1, Appli
14	222	100.0	53	US-09-797-543-5	Sequence 5, Appli
15	222	100.0	53	US-10-016-717-1	Sequence 1, Appli
16	222	100.0	70	US-09-155-076-14	Sequence 14, Appli
17	222	100.0	100	US-09-794-975-4	Sequence 4, Appli
18	222	100.0	103	US-09-972-475-2	Sequence 2, Appli
19	222	100.0	103	US-09-895-443-2	Sequence 2, Appli

20	222	100.0	117	10	US-09-794-975-6	Sequence 6, Appli
21	222	100.0	117	10	US-09-823-153-2	Sequence 2, Appli
22	222	100.0	355	10	US-09-794-975-13	Sequence 13, Appli
23	222	100.0	695	10	US-09-794-927-10	Sequence 10, Appli
24	222	100.0	695	10	US-09-794-927-12	Sequence 12, Appli
25	222	100.0	695	10	US-09-794-927-14	Sequence 14, Appli
26	222	100.0	695	10	US-09-795-847-10	Sequence 10, Appli
27	222	100.0	695	10	US-09-795-847-12	Sequence 12, Appli
28	222	100.0	695	10	US-09-795-847-14	Sequence 14, Appli
29	222	100.0	695	10	US-09-794-743-10	Sequence 10, Appli
30	222	100.0	695	10	US-09-794-743-12	Sequence 12, Appli
31	222	100.0	695	10	US-09-794-743-14	Sequence 14, Appli
32	222	100.0	695	10	US-09-794-748-10	Sequence 10, Appli
33	222	100.0	695	10	US-09-794-748-12	Sequence 12, Appli
34	222	100.0	695	10	US-09-794-748-14	Sequence 14, Appli
35	222	100.0	695	10	US-09-794-925-10	Sequence 10, Appli
36	222	100.0	695	10	US-09-794-925-12	Sequence 12, Appli
37	222	100.0	695	10	US-09-794-925-14	Sequence 14, Appli
38	222	100.0	695	10	US-09-681-442-10	Sequence 10, Appli
39	222	100.0	695	10	US-09-681-442-12	Sequence 12, Appli
40	222	100.0	695	10	US-09-681-442-14	Sequence 14, Appli
41	222	100.0	695	10	US-09-149-718-2	Sequence 2, Appli
42	222	100.0	697	10	US-09-794-927-16	Sequence 16, Appli
43	222	100.0	697	10	US-09-794-927-18	Sequence 18, Appli
44	222	100.0	697	10	US-09-794-927-20	Sequence 20, Appli
45	222	100.0	697	10	US-09-795-847-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-10-076-708-7
; Sequence 7, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Shatma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076.708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-7

Query Match 100.0%; Score 222; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWIAT 43
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGWIAT 43

RESULT 2

US-10-051-496-1
; Sequence 1, Application US/10051496
; Publication No. US20020182660A1
; GENERAL INFORMATION:
; APPLICANT: Kei-Lai L. Fong
; TITLE OF INVENTION: N- and C-Terminus Specific Immunoassays for
; Full Length Beta-Amyloid Peptide - Abeta(1-40), Abeta(1-42), Abeta(1-43)

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: Kei-Lai L. Fong

STREET: 1004 West 8th Avenue

CITY: King of Prussia

STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch, 1.44MB storage
COMPUTER: IBM PC Compatibles
OPERATING SYSTEM: Windows
SOFTWARE: MS No. US20020182660Alepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/051,496
FILING DATE: 18-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/784,854A
FILING DATE: 16-Feb-2001
APPLICATION NUMBER: 60/183,407
FILING DATE: 18-February-2000
ATTORNEY/AGENT INFORMATION:
NAME: Koenig, C. Frederick III
REGISTRATION NUMBER: 29,662
REFERENCE/DOCKET NUMBER: PBI-PT001.1
TELEPHONE: (215) 568-6400
TELEFAX: (215) 568-6499
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 Amino Acid
TYPE: Amino Acid
STRANDEDNESS: <Unknown>
TOPOLOGY: Linear
MOLECULE TYPE: Protein
HYPOTHETICAL: <Unknown>
ANTI-SENSE: <Unknown>
ORIGINAL SOURCE:
ORGANISM: <Unknown>
INDIVIDUAL ISOLATE: <Unknown>
CELL TYPE: <Unknown>
IMMEDIATE SOURCE:
LIBRARY: <Unknown>
CLONE: <Unknown>
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 43
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-051-496-1

Query Match 100.0%; Score 222; DB 9; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43
Db 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43

RESULT 3

US-09-280-966-1
Sequence 1, Application US/09280966
Patent No. US20010020097A1
GENERAL INFORMATION:
APPLICANT: JAMES E. AUDIA
BEVERLY K. FOLMER
VARGHESE JOHN
JEFFREY S. NISSEN
WARREN J. PORTER
EUGENE D. THORSETT
JING WU

TITLE OF INVENTION: N-(ARYL/HETEROARYLACETYL) AMINO
ACID ESTERS, PHARMACEUTICAL COMPOSITIONS
COMPRISING SAME, AND METHODS FOR INHIBITING
-AMYLOID PEPTIDE RELEASE AND/OR ITS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis, LLP
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,966
FILING DATE: 30-Mar-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/976,191
FILING DATE: 21 NOV 1997
APPLICATION NUMBER: 60/077,175
FILING DATE: 22 NOV 1996
ATTORNEY/AGENT INFORMATION:
NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 002010-335
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-622-2300
TELEFAX: 650-622-2499
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-280-966-1
Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43
Db 1 DAEFRHDSGYEVHHQKLVFAEDVGSNKGALIGLMVGGVVIAT 43

RESULT 4

US-09-904-987-1
Sequence 1, Application US/09904987
Patent No. US20020037908A1
GENERAL INFORMATION:
APPLICANT: No. US20020037908Alactyl, Inc.
TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
TITLE OF INVENTION: Protein Assembly or Aggregation
FILE REFERENCE: 42108/26146
CURRENT APPLICATION NUMBER: US/09/904,987
CURRENT FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 43
TYPE: PRT
ORGANISM: homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NCBI ENTREZ / QRHUA4
DATABASE ENTRY DATE: 2000-09-15
RELEVANT RESIDUES: (672)..(714)
US-09-904-987-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 5
US-09-808-037-3
; Sequence 3, Application US/09808037
; Patent No. US20020052311A1
; GENERAL INFORMATION:
; APPLICANT: SOLOMON, Beke
; APPLICANT: HANAN, Eliat
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
; FILE OF INVENTION: NEUROLOGICAL DISEASES AND DISORDERS
; FILE REFERENCE: SOLOMON-2D
; CURRENT APPLICATION NUMBER: US/09/808,037
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/629,971
; PRIOR FILING DATE: 2000-07-31
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/152,417
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-808-037-3

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 6
US-09-866-712-3
; Sequence 3, Application US/09866712
; Patent No. US20020058637A1
; GENERAL INFORMATION:
; APPLICANT: Akhiko TAKASHIMA et al.
; TITLE OF INVENTION: METHODS OF INHIBITING TAU-PROTEIN KINASE I ACTIVITY
; INHIBITING NEURONAL CELL DEATH AND TREATING ALZHEIMER'S DIS
; KINASE I (AS AMENDED)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WENDEROTH, LIND & PONACK
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/866,712
; FILING DATE: 30-May-2001
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/216,958
FILING DATE: December 21, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 2001-0488/LC/00177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)721-8200
TELEFAX: (202)721-8250
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-866-712-3

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 7
US-09-972-475-1
; Sequence 1, Application US/09972475
; Patent No. US20020098173A1
; GENERAL INFORMATION:
; APPLICANT: Findeis, Mark A. et al.
; TITLE OF INVENTION: Modulators of Amyloid Aggregation
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/972,475
; FILING DATE: 04-Oct-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,267
; FILING DATE: <Unknown>
; APPLICATION NUMBER: USSN 08/475,579
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: USSN 08/548,998
; FILING DATE: 27-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PPI-002CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

; Patent No. US20020136718A1
; GENERAL INFORMATION:
; APPLICANT: Raso, Victor
; TITLE OF INVENTION: IMMUNOLOGICAL CONTROL OF BETA-AMYLOID LEVELS IN VIVO
; FILE REFERENCE: BBRI-2005
; CURRENT APPLICATION NUMBER: US/09/992,994
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/594,366
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,408
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-994-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 12
US-09-984-834-1
; Sequence 1, Application US/09984834
; Patent No. US20020137743A1
; GENERAL INFORMATION:
; APPLICANT: JING WU
; EUGENE D. THORSETT
; JEFFREY S. NISSEN
; THOMAS E. MABRY
; LEE H. LATIMER
; VARGHESE JOHN
; LAWRENCE Y. FANG
; JAMES E. AUDIA
; TITLE OF INVENTION: N-(ARYL/HERETOARYLACETYL)
; AMINO ACID ESTERS, PHARMACEUTICAL
; COMPOSITIONS COMPRISING SAME, AND
; METHODS FOR INHIBITING BETA-AMYLOID
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker &
; Mathis, LLP
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,834
; FILING DATE: 31-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/976,179
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Swiss, Gerald F.
; REGISTRATION NUMBER: 30,113
; REFERENCE/DOCKET NUMBER: 002010-050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-622-2300

; TELEFAX: 650-622-2499
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-984-834-1

Query Match 100.0%; Score 222; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 13
US-10-041-605-1
; Sequence 1, Application US/10041605
; Patent No. US20020115223A1
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bush, Rudolph I.
; TITLE OF INVENTION: An In Vitro System For Determining Formation of A Amyloid
; FILE REFERENCE: 0609.4100001
; CURRENT APPLICATION NUMBER: US/10/041,605
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 08/294,819
; PRIOR FILING DATE: 1994-08-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-605-1

Query Match 100.0%; Score 222; DB 12; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

RESULT 14
US-09-797-543-5
; Sequence 5, Application US/09797543
; Patent No. US20020072050A1
; GENERAL INFORMATION:
; APPLICANT: Hook, Vivian Y.H.
; TITLE OF INVENTION: SECRETASES RELATED TO ALZHEIMER'S DEMENTIA
; FILE REFERENCE: P-AS 4579
; CURRENT APPLICATION NUMBER: US/09/797,543
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 09/173,887
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 53
; TYPE: PRT
; ORGANISM: 'Axial Seamount' polynoid polychaete
US-09-797-543-5

Query Match 100.0%; Score 222; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 Db 4 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 46

RESULT 15

US-10-016-717-1
 ; Sequence 1, Application US/10016717
 ; Patent No. US20020132281A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hook, Vivian Y.H.
 ; TITLE OF INVENTION: SECRETSSES RELATED TO ALZHEIMER'S DEMENTIA
 ; FILE REFERENCE: P-AS 5031
 ; CURRENT APPLICATION NUMBER: US/10/016.717
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: US 09/173,887
 ; PRIOR FILING DATE: 1998-10-16
 ; PRIOR APPLICATION NUMBER: US 09/294,987
 ; PRIOR FILING DATE: 1999-04-20
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 53
 ; TYPE: PRT
 ; ORGANISM: mammalian
 US-10-016-717-1

Query Match 100.0%; Score 222; DB 12; Length 53;
 Best Local Similarity 100.0%; Pred. No. 2e-23;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 43
 Db 4 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT 46

Search completed: January 28, 2003, 14:40:03
 Job time : 1.56378 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 0.892652 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-1

Perfect score: 222

Sequence: 1 DAEFRHDSGEVHHQKLVFF.....VGSNKGAIGLMVGGVVIAT 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222	100.0	43	1	US-08-235-400-1
2	222	100.0	43	1	US-08-437-067-1
3	222	100.0	43	1	US-08-302-808-6
4	222	100.0	43	1	US-08-079-511-1
5	222	100.0	43	1	US-08-467-607-1
6	222	100.0	43	2	US-08-404-831-1
7	222	100.0	43	2	US-08-602-264A-3
8	222	100.0	43	2	US-08-469-362-1
9	222	100.0	43	2	US-08-612-785B-1
10	222	100.0	43	2	US-08-475-579A-1
11	222	100.0	43	2	US-08-850-392-1
12	222	100.0	43	2	US-08-986-348-6
13	222	100.0	43	2	US-08-975-977-1
14	222	100.0	43	2	US-08-817-423-1
15	222	100.0	43	2	US-08-920-162A-1
16	222	100.0	43	3	US-08-451-018A-3
17	222	100.0	43	3	US-08-976-191-1
18	222	100.0	43	3	US-08-976-179-1
19	222	100.0	43	4	US-09-216-958-3
20	222	100.0	43	4	US-09-356-931-1
21	222	100.0	43	4	US-08-733-202-1
22	222	100.0	43	4	US-08-703-675C-1
23	222	100.0	43	4	US-09-390-692-1
24	222	100.0	43	4	US-08-617-267C-1
25	222	100.0	43	4	US-09-303-655-1
26	222	100.0	43	4	US-08-294-819-1
27	222	100.0	47	2	US-08-609-090-10

28	222	100.0	52	2	US-08-609-090-11	Sequence 11, Appl
29	222	100.0	53	4	US-09-173-887-5	Sequence 5, Appl
30	222	100.0	59	1	US-08-484-969-3	Sequence 3, Appl
31	222	100.0	59	1	US-08-472-627-3	Sequence 3, Appl
32	222	100.0	59	1	US-08-388-463-3	Sequence 3, Appl
33	222	100.0	63	1	US-08-462-859A-4	Sequence 4, Appl
34	222	100.0	63	1	US-08-123-659A-4	Sequence 4, Appl
35	222	100.0	63	1	US-08-464-247A-4	Sequence 4, Appl
36	222	100.0	63	1	US-08-464-248A-4	Sequence 4, Appl
37	222	100.0	99	3	US-08-422-333-3	Sequence 3, Appl
38	222	100.0	99	3	US-08-339-708A-4	Sequence 4, Appl
39	222	100.0	99	3	US-08-339-708A-4	Sequence 4, Appl
40	222	100.0	100	6	5187153-10	Sequence 8, Appl
41	222	100.0	100	6	520013-10	Patent No. 5187153
42	222	100.0	100	6	523482-10	Patent No. 523482
43	222	100.0	103	2	US-08-404-831-2	Sequence 2, Appl
44	222	100.0	103	2	US-08-612-785B-2	Sequence 2, Appl
45	222	100.0	103	2	US-08-475-579A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-08-235-400-1

Sequence 1, Application US/08235400

Patent No. 5552426

GENERAL INFORMATION:

APPLICANT: Lunn, William H.

APPLICANT: Monn, James A.

APPLICANT: Zimmerman, Dennis M.

TITLE OF INVENTION: METHODS FOR TREATING A PHYSIOLOGICAL

DISORDER ASSOCIATED WITH BETA AMYLOID PEPTIDE

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center/1104

CITY: Indianapolis

STATE: Indiana

COUNTRY: United States of America

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/235,400

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Gaylo, Paul J.

REGISTRATION NUMBER: 36,808

REFERENCE/DOCKET NUMBER: X-9507

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 276-0756

TELEFAX: (317) 276-3861

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 43 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-235-400-1

Query Match 100.0%; Score 222; DB 1; Length 43;

Best Local Similarity 100.0%; Pred. No. 3.9e-27;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGEVHHQKLVFFAEEDVGSNKGAIGLMVGGVVIAT 43

Db 1 DAEFRHDSGEVHHQKLVFFAEEDVGSNKGAIGLMVGGVVIAT 43

RESULT 2
US-08-437-067-1
; Sequence 1, Application US/08437067
; Patent No. 5593846
; GENERAL INFORMATION:
; APPLICANT: Schenk, Dale B.
; APPLICANT: Selkoe, Dennis J.
; APPLICANT: Schlossmacher, Michael G.
; APPLICANT: Seubert, Peter A.
; APPLICANT: Vigo-Pelfrey, Carmen
; TITLE OF INVENTION: Methods and Compositions for
; TITLE OF INVENTION: Detection
; TITLE OF INVENTION: of Soluble Beta-Amyloid Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/437,067
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,972
; FILING DATE: 26-OCT-1992
; APPLICATION NUMBER: US 07/911,647
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15270-6-1
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-437-067-1
Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 3
US-08-302-808-6
; Sequence 6, Application US/08302808
; Patent No. 5750349
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, No. 5750349uhiro
; APPLICANT: ODAKA, Asano
; APPLICANT: KITADA, Chieko
; TITLE OF INVENTION: ANTIBODIES TO B-AMYLOIDS OR THEIR
; TITLE OF INVENTION: DERIVATIVES AND USE THEREOF
; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/302,808
; FILING DATE: 15-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00089
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 010132/1993
; FILING DATE: 25-JAN-1993
; APPLICATION NUMBER: 019035/1993
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 286985/1993
; FILING DATE: 16-NOV-1993
; APPLICATION NUMBER: 334773/1993
; FILING DATE: 28-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 44631
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-302-808-6
Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
RESULT 4
US-08-079-511-1
; Sequence 1, Application US/08079511
; Patent No. 5766846
; GENERAL INFORMATION:
; APPLICANT: Schenk, Dale B.
; APPLICANT: Selkoe, Dennis J.
; APPLICANT: Schlossmacher, Michael G.
; APPLICANT: Seubert, Peter A.
; APPLICANT: Vigo-Pelfrey, Carmen
; TITLE OF INVENTION: Methods and Compositions for Detection
; TITLE OF INVENTION: of Soluble Beta-Amyloid Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Tower, Suite 2000

City: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/079,511
FILING DATE: 19930617
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,972
FILING DATE: 26-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15270-6-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-079-511-1

Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 5

US-08-467-607-1
Sequence 1, Application US/08467607
Patent No. 5783434
GENERAL INFORMATION:

APPLICANT: TUNG, JAY S.
APPLICANT: SINHA, SUKANTO
APPLICANT: MCCONLOGUE, LISA
APPLICANT: TATSUNO, GWEN
APPLICANT: ANDERSON, JOHN
APPLICANT: CHRYSLER, SUSANNA
TITLE OF INVENTION: NOVEL CATHESPIN AND METHODS AND
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: ATHENA NEUROSCIENCES
STREET: 800 F. Gateway Blvd.
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,607
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DUVALL, JEAN M.

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

REGISTRATION NUMBER: 32,731
REFERENCE/DOCKET NUMBER: 002010-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-607-1

Query Match 100.0%; Score 222; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
DB 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 6

US-08-404-831-1
Sequence 1, Application US/08404831
Patent No. 5817626
GENERAL INFORMATION:

APPLICANT: Mark A. Findeis, Howard Benjamin, Marc B. Garnick,
APPLICANT: Malcolm L. Geftter, Arvind Hundal, Laura Kasman,
APPLICANT: Gary Musso, Ethan R. Signer, and James Wakefield
TITLE OF INVENTION: Modulators of {SYMBOL 98 \ "Symbol"}-Amyloid Peptide Aggre
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,831
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hanley, Elizabeth A. (EAH)
REGISTRATION NUMBER: 33,505
REFERENCE/DOCKET NUMBER: PPI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-08-404-831-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

US-08-602-264A-3
Sequence 3, Application US/08602264A
Patent No. 5837853
GENERAL INFORMATION:
APPLICANT: AKIHiko TAKASHIMA et al.
TITLE OF INVENTION: PREVENTIVE OR THERAPEUTIC AGENTS FOR
TITLE OF INVENTION: ARZHEIMER'S DISEASE, A SCREENING METHOD OF ARZHEIMER'S DISEASE
TITLE OF INVENTION: tau-PROTEIN KINASE I ORIGINATED FROM HUMAN BEING (AS AMENDED)
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: WENDEROTH, LIND & PONACK
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
COUNTRY: D.C.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch,
MEDIUM TYPE: 144 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602.264A
FILING DATE: February 20, 1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,091
FILING DATE: March 2, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-602-264A-3

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 8
US-08-469-362-1
Sequence 1, Application US/08469362
Patent No. 5849711
GENERAL INFORMATION:
APPLICANT: TUNG, JAY S.
APPLICANT: SINHA, SUKANTO
APPLICANT: MCCONLOGUE, LISA
APPLICANT: SEMKO, CHRISTOPHER M.F.
TITLE OF INVENTION: NOVEL CATHEPSIN AND METHODS AND
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: ATHENA NEUROSCIENCES
STREET: 900 F. Gateway Blvd.

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 9
US-08-612-785B-1
Sequence 1, Application US/08612785B
Patent No. 5854204
GENERAL INFORMATION:
APPLICANT: FIndels, Mark A. et al.
TITLE OF INVENTION: Ab Peptides that Modulate b-Amyloid
TITLE OF INVENTION: Aggregation
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,785B
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/404,831
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/475,579
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/548,998
FILING DATE: 27-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A.

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 9
US-08-612-785B-1
Sequence 1, Application US/08612785B
Patent No. 5854204
GENERAL INFORMATION:
APPLICANT: FIndels, Mark A. et al.
TITLE OF INVENTION: Ab Peptides that Modulate b-Amyloid
TITLE OF INVENTION: Aggregation
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,785B
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/404,831
FILING DATE: 14-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/475,579
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/548,998
FILING DATE: 27-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A.

REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PPI-002CP3
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-612-785B-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43

RESULT 10
US-08-475-579A-1
Sequence 1, Application US/08475579A
Patent No. 5854215
GENERAL INFORMATION:

APPLICANT: Mark A. Findels et al.
TITLE OF INVENTION: Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Peptide Aggrega
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,579A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,831
FILING DATE: 14-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P41,106
REFERENCE/DOCKET NUMBER: PPI-002CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-475-579A-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43

RESULT 11

US-08-850-392-1

Sequence 1, Application US/08850392
Patent No. 585982
GENERAL INFORMATION:
APPLICANT: TUNG, JAY S.
APPLICANT: SINHA, SURANTO
APPLICANT: MCCONLOGUE, LISA
APPLICANT: SEMKO, CHRISTOPHER M.F.
TITLE OF INVENTION: NOVEL CATHPSIN AND METHODS AND
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: ATHENA NEUROSCIENCES
STREET: 800 F. Gateway Blvd.
CITY: South San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,392
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/469,362
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: DUVALL, JEAN M.
REGISTRATION NUMBER: 32,731
REFERENCE/DOCKET NUMBER: 002010-005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-850-392-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||

Db 1 DAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIAT 43
|||||

RESULT 12

US-08-986-948-6

Sequence 6, Application US/08986948
Patent No. 5955317
GENERAL INFORMATION:
APPLICANT: SUZUKI, No. 5955317uhiro
APPLICANT: ODAKA, Asano
APPLICANT: KITADA, Chieko
TITLE OF INVENTION: ANTIBODIES TO B-AMYLOIDS OR THEIR
TITLE OF INVENTION: DERIVATIVES AND USE THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON

RESULT 13
US-08-975-977-1
; Sequence 1, Application US/08975977
; Patent No. 5965614
; GENERAL INFORMATION:
; APPLICANT: JAMES E. AUDIA
; APPLICANT: BEVERLY K. FOLMER
; APPLICANT: VARGHESE JOHN
; APPLICANT: LEE H. LATIMER
; APPLICANT: JEFFREY S. NISSEN
; APPLICANT: JON K. REEL
; APPLICANT: EUGENE D. THORSETT
; APPLICANT: CELIA A. WHITESITT
; TITLE ~~OF~~ INVENTION: N-(ARYL/HETEROARYL) AMINO
; ACID ESTERS, PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITIONS COMPRISING SAME, AND
; TITLE OF INVENTION: METHODS FOR INHIBITING BETA-AMYLOID
; TITLE OF INVENTION: PEPTIDE RELEASE AND/OR ITS

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/11895
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0609.4110000/REF/JUK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-817-423-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

RESULT 15

US-08-920-162A-1
Sequence 1, Application US/08920162A
Patent No. 5985242
GENERAL INFORMATION:

APPLICANT: Findeis, M. et al.
TITLE OF INVENTION: Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Peptide
TITLE OF INVENTION: Aggregation Comprising D-Amino Acids
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,162A
FILING DATE: Herewith
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/548,998
FILING DATE: 27-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/616,081
FILING DATE: 14-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/703,675
FILING DATE: 27-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 21-JUL-1997

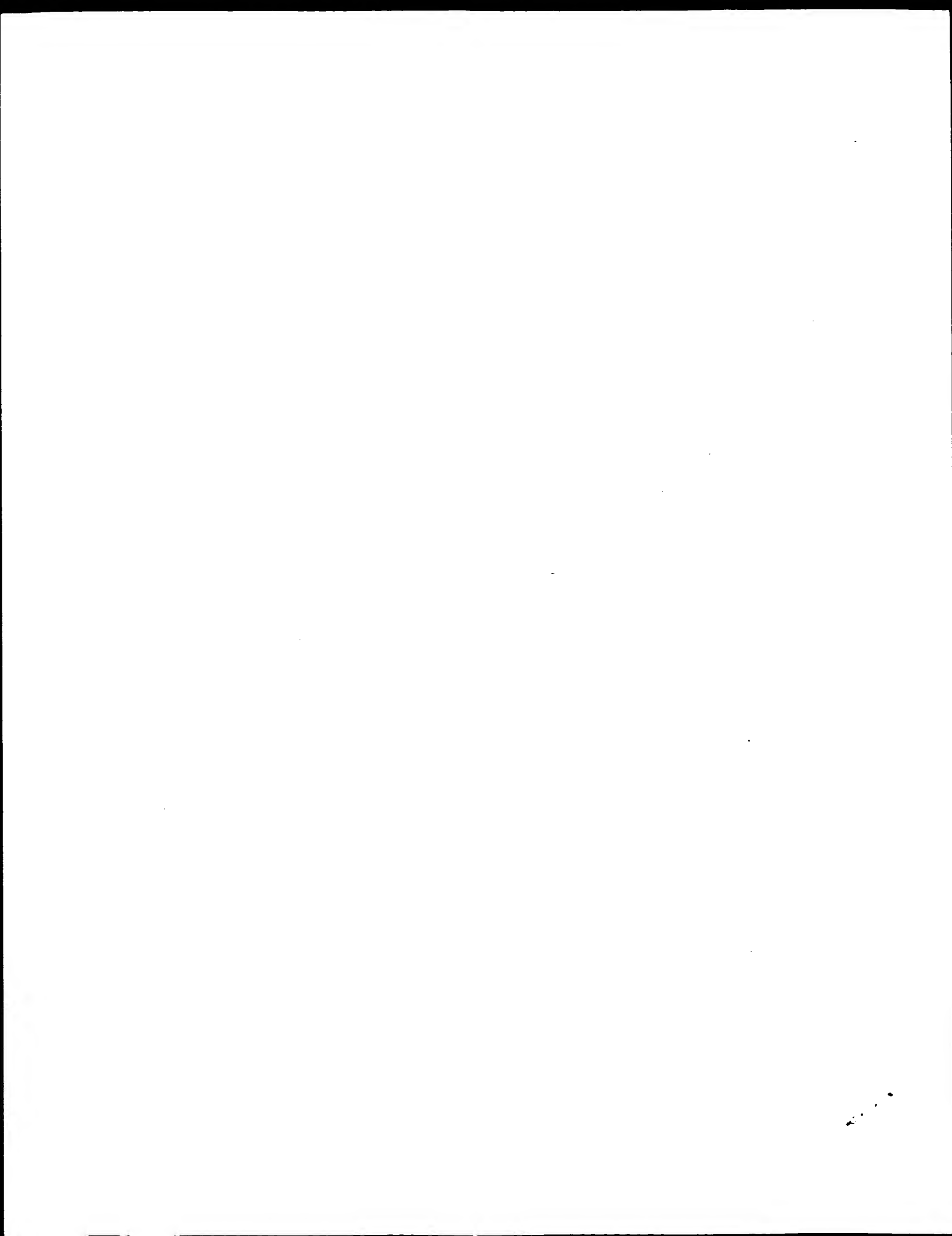
ATTORNEY/AGENT INFORMATION:
NAME: KARA, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: PPI-016CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-920-162A-1

Query Match 100.0%; Score 222; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43
|||||
Db 1 DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVAT 43

Search completed: January 28, 2003, 14:21:17
Job time : 2.89265 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 20.1912 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-2
Perfect score: 4058
Sequence: 1 MLPGLALLLLAAWTAARALEV.....KMQQNGYENPTYKFFEQMQN 770
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	4058	100.0	770	1	QRHUA4	1	Alzheimer's diseases
2	3590.5	88.5	695	1	A49795	2	Alzheimer's diseases
3	3493.5	86.1	695	2	S00550	2	Alzheimer's diseases
4	3468.5	85.5	695	2	A27485	2	Alzheimer's diseases
5	3403.5	83.9	747	2	JH0773	2	Alzheimer's diseases
6	2446.5	60.3	484	4	A32761	4	hypothetical Alzhe
7	1981.5	48.8	763	2	A49321	2	amyloid beta (A4)
8	1971.5	48.6	765	2	S42880	2	amyloid precursor-
9	1956.5	48.2	751	2	A49974	2	beta-amyloid precu
10	1150.5	28.4	653	2	A46362	2	amyloid precursor
11	1112.5	27.4	511	2	JC1404	2	CDEI-box DNA-bindi
12	785	19.3	686	2	T15795	2	hypothetical prote
13	737.5	18.2	886	2	A32758	2	beta-amyloid-like
14	706	17.4	246	2	S38344	2	CDEI-binding prote
15	501	12.3	100	2	A32282	2	Alzheimer's diseases
16	411	10.1	82	2	P00438	2	Alzheimer's diseases
17	403	9.9	76	2	S06678	2	Alzheimer's diseases
18	397	9.8	76	2	S03607	2	Alzheimer's diseases
19	383	9.4	76	2	S04855	2	Alzheimer's diseases
20	296.5	7.3	191	2	A35981	2	sperm membrane pro
21	283	7.0	57	2	P60045	2	Alzheimer's diseases
22	283	7.0	57	2	F60045	2	Alzheimer's diseases
23	283	7.0	57	2	D60045	2	Alzheimer's diseases
24	283	7.0	57	2	D60045	2	Alzheimer's diseases
25	283	7.0	57	2	A60045	2	Alzheimer's diseases
26	283	7.0	57	2	B60045	2	Alzheimer's diseases
27	256.5	6.3	111	2	S41082	2	amyloid precursor
28	217	5.3	42	2	PN0512	2	beta-amyloid prote
29	194.5	4.8	2225	2	T26063	2	hypothetical prote

ALIGNMENTS

RESULT 1

QRHUA4

N:Alzheimer's disease amyloid beta protein precursor [validated] - human

N:Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIA inh
N:Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascu
protein precursor splice form APP(770)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000

C:Accession: S02260; S05194; A32277; A32260; A35486; I39452; I39451; I39453; I59562;

4668; A28583; A29302; A60805; J10038; S06121; A60355; A59011; A38384; S28076; S38252;

R:Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.;

Nucleic Acids Res. 17, 517-522, 1989

A>Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encode

A:Reference number: S02260; MUID:89128427; PMID:2783775

A:Accession: S02260

A:Molecule type: DNA

A:Residues: 1-288, 'V', 365-770 <LEM1>

A:Cross-references: EMBL:X13466

A>Note: alternative splice form APP(695)

R:Lemaire, H.G.

submitted to the EMBL Data Library, November 1988

A:Reference number: S05194

A:Accession: S05194

A:Molecule type: DNA

A:Residues: 1-14, 'VW', 17-288, 'V', 365-770 <LEM2>

A:Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360

A>Note: alternative splice form APP(695)

R:La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.

Biochem. Biophys. Res. Commun. 159, 297-304, 1989

A>Title: Characterization of the 5'-end region and the first two exons of the beta-pr

A:Reference number: A32277; MUID:89165870; PMID:2538123

A:Accession: A32277

A:Molecule type: DNA

A:Residues: 1-75 <LAF>

A:Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AA013654.1; PID:g516074

R:Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.

Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989

A>Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows simila

A:Reference number: A33260; MUID:89392030; PMID:2675837

A:Accession: A33260

A:Molecule type: DNA

A:Residues: 656-737 <JOH>

A:Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865

R:Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B

Biochem. Biophys. Res. Commun. 170, 301-307, 1990

A>Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid

A:Reference number: A35486; MUID:90321244; PMID:2196678

A:Accession: A35486

A:Molecule type: DNA

A:Residues: 672-710 <PRE1>

A>Note: 693-Gln was found in DNA isolated from HCHWA-D patients

R:Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.

Gene 87, 257-263, 1990
A:Title: Genomic organization of the human amyloid beta-protein precursor gene.
A:Reference number: I39451; MUID:90236318; PMID:2110105
A:Accession: I39452
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: DNA
A:Residues: 1-770 <YOS1>
A:Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A:Accession: I39451
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
A:Molecule type: DNA
A:Residues: 1-530, 'OQLMPVTPAFWEAKVGR' <YOS2>
A:Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R:Yoshikawa, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A:Reference number: A59020; MUID:91340168; PMID:1908403
A:Contents: annotation: erratum
A>Note: revised physical map for reference I39451
R:Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duine
Science 248, 1124-1126, 1990
A:Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage
A:Reference number: I39453; MUID:90260663; PMID:2111584
A:Accession: I39453
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 656-737 <LEV>
A:Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A>Note: a mutation with 693-Gln is presented
R:Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A:Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer
A:Reference number: I59562; MUID:92022553; PMID:1925564
A:Accession: I59562
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 689-716, 'F', 718-737 <MUR>
A:Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R:Ramano, K.; Orr, H.T.; Payami, E.M.; Alonso, M.E.; Pulst, S.M.; Anderson, R.;
arakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin,
Am. J. Hum. Genet. 51, 998-1014, 1992
A:Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the
A:Reference number: A44017; MUID:93035397; PMID:1415269
A:Accession: A44017
A:Molecule type: DNA
A:Residues: 687-718 <KAN2>
A:Cross-references: GB:S45136; NID:g257377; PIDN:AAB23645.1; PID:g257378
A:Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A:Experimental source: familial Alzheimer disease family SB
A>Note: sequence extracted from NCBI backbone (NCBIP:115374)
A:Accession: B44017
A:Molecule type: DNA
A:Residues: 687-718 <KAN2>
A:Cross-references: GB:S45136; NID:g257377; PIDN:AAB23646.1; PID:g257380
A:Experimental source: familial Alzheimer disease family LIT
A>Note: sequence extracted from NCBI backbone (NCBIP:115376)
A>Note: this sequence has a silent mutation
R:Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.;
Nature 325, 733-736, 1987
A:Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface
A:Reference number: A03134; MUID:87144572; PMID:2881207
A:Accession: A03134
A:Molecule type: mRNA
A:Residues: 1-288, 'V', 365-770 <KAN>
A:Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A>Note: alternative splice form APP(695)
R:Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A:Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular A
A:Reference number: A29030; MUID:87231971; PMID:3035574
A:Accession: A29030
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-646, 'E', 648-770 <ROB>
A:Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A>Note: the authors translated the codon GAG for residue 647 as Asp

R:Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A:Title: Characterization and chromosomal localization of a cDNA encoding brain amylo
A:Reference number: A47584; MUID:87120328; PMID:3810169
A:Accession: A47584
A:Molecule type: mRNA
A:Residues: 674-756, 'S', 758-770 <GOL>
A:Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
A:Experimental source: brain
R:Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van
Science 235, 880-884, 1987
A:Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near
A:Reference number: A47585; MUID:87120329; PMID:2949367
A:Accession: A47585
A:Molecule type: mRNA
A:Residues: 674-703 <TANI>
A:Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
R:Dykes, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mue
EMBO J. 7, 949-957, 1988
A:Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 p
A:Reference number: S02638; MUID:88296437; PMID:2900137
A:Accession: S02638
A:Molecule type: mRNA
A:Residues: 672-678 <DYR>
R:Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; N
Nature 331, 528-530, 1988
A:Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA assoc
A:Reference number: S00707; MUID:88122640; PMID:2893290
A:Accession: S00707
A:Molecule type: mRNA
A:Residues: 286-344, 'I', 365-366 <TAN2>
A:Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g292612
A:Experimental source: promyelocytic leukemia cell line HL60
A>Note: alternative splice form APP(751)
R:Ponte, P.; Gonzalez-Dewhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.;
Nature 331, 535-537, 1988
A:Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inh
A:Reference number: S00925; MUID:88122639; PMID:2893289
A:Accession: S00925
A:Molecule type: mRNA
A:Residues: 1-344, 'I', 365-770 <PO>
A:Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A>Note: alternative splice form APP(751)
R:Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A:Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibi
A:Reference number: A38949; MUID:88122641; PMID:2893291
A:Accession: A38949
A:Molecule type: mRNA
A:Residues: 287-367 <KIT>
A:Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g292611
A:Experimental source: glioblastoma cell line
A>Note: alternative splice form APP(770)
R:Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ash
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A:Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three
A:Reference number: A30320
A:Accession: A30320
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 284-288, 'V', 365-770 <VIT1>
A:Accession: B30320
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 122-288, 'V', 365-770 <VIT2>
A:Accession: C30320
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 606-770 <VIT3>
R:Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta,
Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
A:Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease
A:Reference number: A31087; MUID:88124954; PMID:2893379

A:Accession: A31087
A:Molecule type: mRNA
A:Residues: 507-770 <ZAI>
A:Cross-references: GB:M18734; NID:g178572; PIDN:AAAS1726.1; PID:g178573
A>Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603
8 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 65
A>Note: the cited Genbank accession number, J03594, is not in release 101.0
R:Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.4e-210;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRLNMHMNVQNGKWDSDPSGK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRLNMHMNVQNGKWDSDPSGK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPVEATERTTSTATTTTTTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPVEATERTTSTATTTTTTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEP LARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEP LARD 360
QY 361 PVKLPTTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 361 PVKLPTTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWDPKKAQIRSOVMTHLRVIER 540
Db 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWDPKKAQIRSOVMTHLRVIER 540
QY 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
Db 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
QY 721 VMLKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEOMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEOMQN 770

RESULT 2
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A49795
R:Podlisky, M.B.; Tolan, D.R.; Selkoe, D.J.

Am. J. Pathol. 138, 1423-1435, 1991
A:Title: Homology of the amyloid beta protein precursor in monkey and human supports
A:Reference number: A49795; MUID:91273117; PMID:1905108
A:Accession: A49795.
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-695 <POD>
A:Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C:Superfamily: Alzheimer's disease amyloid beta protein; animal kunitz-type proteinases
C:Keywords: alternative splicing

Query Match 88.5%; Score 3590.5; DB 1; Length 695;
Best Local Similarity 90.1%; Pred. No. 5.3e-185;
Matches 694; Conservative 1; Mismatches 0; Indels 75; Gaps 1;
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRLNMHMNVQNGKWDSDPSGK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRLNMHMNVQNGKWDSDPSGK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPVEATERTTSTATTTTTTESVEEVVREVCSQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPVEATERTTSTATTTTTTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEP LARD 360
Db 289 ----- 288
QY 361 PVKLPTTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
Db 289 ---VPTTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 345
QY 421 KNLPKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 346 KNLPKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWDPKKAQIRSOVMTHLRVIER 540
Db 406 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRWDPKKAQIRSOVMTHLRVIER 465
QY 541 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
Db 466 MNQSLLYNPVPAVEEIQDEVELLQKEQNYSDVLANMISEPRISYGNALMPSLTET 525
QY 601 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 526 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 645
QY 721 VMLKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEOMQN 770
Db 646 VMLKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEOMQN 695

RESULT 3
S00550
Alzheimer's disease amyloid beta protein precursor - rat
N:Alternate names: beta-A4 amyloid protein
C:Species: Rattus norvegicus (Norway rat)

[illegible]

	Query Match	85.5%	Score 3468.5	DB 2	Length 695
	Best Local Similarity	87.3%	Pred. No. 1.8e-178		
	Matches 672	Conservative	6	Mismatches 17	Indels 75
	Gaps	1			
Qy	1	MLPGLALLLLAAATW	RALEVPTDGNAGLLAEPQ	AMFCGRLLNMHMNVQNGKWDSDPGTK	60
Db	1	MLPSALLLLAAATW	RALEVPTDGNAGLLAEPQ	AMFCGLNMHMNVQNGKWDSDPGTK	60
Qy	61	TCIDTKEGILQYCEVY	PELQITNVVEANQPVTTQ	NCKRGRKCKCTHPHFVYPRCLVG	120
Db	61	TCIGTKEGILQYCEVY	PELQITNVVEANQPVTTQ	NCKRGRKCKCTHTHIVYPRCLVG	120
Qy	121	EFVSDALLVDPKCKFL	HQERMDVCGETHLHHHTV	AKETCSKSTNLHDYGMGLPCGIDKFR	180

A:Description: Complete nucleotide ad deduced amino acid sequence of rat amyloid precursor
A:Reference number: S42880
A:Accession: S42880
A:Molecule type: mRNA
A:Residues: 1-765 <SA2>
A:Cross-references: EMBL:X77934
R:Sandbrink, R.; Masters, C.L.; Beyreuther, K.
Biochim. Biophys. Acta 1219, 167-170, 1994
A:Title: Complete nucleotide and deduced amino acid sequence of rat amyloid protein precursor
A:Reference number: S47528; UID:94368849; PMID:8086458
A:Accession: S47528
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-765 <SA2>
A:Cross-references: EMBL:X77934
C:Superfamily: Alzheimer's disease amyloid beta protein; animal kunitz-type proteinase
F:312-362/Domain: animal kunitz-type proteinase inhibitor homology <BPI>
Query Match 48.6%; Score 1971.5; DB 2; Length 765;
Best Local Similarity 49.7%; Pred. No. 3.3e-98;
Matches 401; Conservative 134; Mismatches 173; Indels 99; Gaps 20;
QY 5 LALLLLAAWTAARALEV-----PTDGNAG---LLAEPOIAMFCGRLLNMHNVONGKWDSDP 56
DB 15 LVLLLLGLTAPAAALAGYIEALAAAGTGFVAEPOIAMFCGKLNHNVNIQTGWEDDP 74
QY 57 SGTCTCIDTKEGILQYCOEYVPELQITNVVEANQPTVIONWCKRGRKCKTHPHFVPIYR 116
DB 75 TGTCKSLGTKEEVLYQCOEYVPELQITNVVEANQPTVIONWCKRGRKCKTHPHFVPIYR 132
QY 117 CLVGEFVSALLVPDPCKFLHQRMDVCETHLHWHTVAKETCTSEKSTNLHDYGMLLPCGI 176
DB 133 CLVGEFVSALLVPDPCKFLHQRMDVCETHLHWHTVAKETCTSEKSTNLHDYGMLLPCGI 192
QY 177 DKFRGVFVCCPLAE--ESDNVDSADAEEDSDVWVGADTDYA--DGSEDKVVEVAEE 233
DB 193 DQFHGTGYVCCPQTQVYVDSSTMSKEEEEEE--DEEDYALDKSEFTEADLEDT 248
QY 234 VAEVEEEDDDDEDEDEVEEAEPEE-----ATERTSIATTTTTTSESVEV 287
DB 249 EAAADEDEDEDEDEVEEVEEDRYVDSFKGDDYNEENPTEPSSDGTISDKETAHDY 308
QY 288 REVCSEQAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNFDPEEYCMVCGSAMSO 347
DB 309 KAVCSQEAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNFDPEEYCMVCGSAMSO 366
QY 348 SLLKTTQEPILARDPVLPTTAASTPDVADKYLETPGDENEHAFQKAKERLEAKHRMS 407
DB 367 -----PPTPLPT-----NDVDVYFETSADDNEHAFQKAKERLEAKHRMS 408
QY 408 QVMREWEAEARQAKNLPKADKAVIQHFQKVESLBOEAEANERQQLVETHMARVAMLND 467
DB 409 RVKKEWEAEARQAKNLPKADKAVIQHFQKVESLBOEAEANERQQLVETHMARVAMLND 468
QY 468 RRLALENYITALQAVPPRPRHVFNLKKYVRAEQKDROHTLKHFEHVRMVDPKAAQIR 527
DB 469 RRLALENYITALQAVPPRPRHVFNLKKYVRAEQKDROHTLKHFEHVRMVDPKAAQIR 528
QY 528 SQVMTLRLVYIERMNSLLLYNVPVAEEIQDEVDLQKQNTYSDVLANWISSEPRIS 587
DB 529 SQVMTLRLVYIERMNSLLLYNVPVAEEIQDEVDLQKQNTYSDVLANWISSEPRIS 587
QY 588 YGNDALMPSLTETKTTVELLPVNGEFLSDLOPWHSGADSPANTENEVEPVDARPAD 647
DB 575 ----DQFTSSISENPVDVR---VSSSES--EIPPFPLHPF--HPFPLHPF-----ADP 621
QY 648 RGLTRPGSLTN-----IKTEE--ISEVKMDAERHDSGVEVHOKLVFAEDVGS-- 697
DB 622 HPM--KKGSGMAEQDGLIGAEKVINSKNDENVIDETLDV--KEMIFNAERVGGLE 677
QY 698 -----NKGAIGLMGVGVVIAIVITLVMKKKQYTSIHGVVEVDAV 742
DB 698 -----NKGAIGLMGVGVVIAIVITLVMKKKQYTSIHGVVEVDAV 742

DB 678 EEPDSVGPLRDFSLSSSALIGLLVIAVAIAIVIVISLVMLRKRQYGTISHGIVEHPML 737
QY 743 TPEERHLKMQONGYENPTYKFFEQMQ 769
DB 738 TPEERHLKMQONGYENPTYKFFEQMQ 764
RESULT 9
A49974
beta-amyloid precursor protein 2 homolog APLP2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C:Accession: A49974
R:Slunt, H.H.; Thinakaran, G.; Von Koch, C.; Lo, A.C.; Tanzi, R.E.; Sisodia, S.S.
J. Biol. Chem. 269, 2637-2644, 1994
A:Title: Expression of a ubiquitous, cross-reactive homologue of the mouse beta-amyloid
A:Reference number: A49974; UID:94132029; PMID:8300594
A:Accession: A49974
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-751 <SU>
A:Cross-references: GB:U15571; NID:g558467; PIDN:AAA50603.1; PID:g558468
A:Note: sequence extracted from NCBI backbone (NCBI:144636)
C:Superfamily: Alzheimer's disease amyloid beta protein; animal kunitz-type proteinase
F:310-360/Domain: animal kunitz-type proteinase inhibitor homology <BPI>
Query Match 48.2%; Score 1956.5; DB 2; Length 751;
Best Local Similarity 49.3%; Pred. No. 2e-97;
Matches 397; Conservative 130; Mismatches 169; Indels 109; Gaps 20;
QY 5 LALLLLAAWTAARALEV-----PTDGNAG---LLAEPOIAMFCGRLLNMHNVONGKWDSDP 56
DB 15 LVLLLLGLTAPAAALAGYIEALAAAGTGFVAEPOIAMFCGKLNHNVNIQTGWEDDP 74
QY 57 SGTCTCIDTKEGILQYCOEYVPELQITNVVEANQPTVIONWCKRGRKCKTHPHFVPIYR 116
DB 75 TGTCKSLGTKEEVLYQCOEYVPELQITNVVEANQPTVIONWCKRGRKCKTHPHFVPIYR 132
QY 117 CLVGEFVSALLVPDPCKFLHQRMDVCETHLHWHTVAKETCTSEKSTNLHDYGMLLPCGI 176
DB 133 CLVGEFVSALLVPDPCKFLHQRMDVCETHLHWHTVAKETCTSEKSTNLHDYGMLLPCGI 192
QY 177 DKFRGVFVCCPLAE--ESDNVDSADAEEDSDVWVGADTDYA--DGSEDKVVEVAE--E 231
DB 193 DQFHGTGYVCCPQTQVYVDSSTMSKEEEEEE--DEEDYALDKSEFTEADLEDT 243
QY 232 EVAEVEEEDAD--DEEDDEDEDEVEEAE-----EPYEAETERTSIATTTTTTTES 282
DB 244 ADLEDFTEAADADEDEDEVEEVEEDRYVDSFKGDDYNE--ENPTEPSEGTISDK 301
QY 283 VEEVREVCSQAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECMVCG 342
DB 302 IVHDVKVCSQAETGPCRAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECMVCG 361
QY 343 SAMSLLKTTQEPILARDPVLPTTAASTPDVADKYLETPGDENEHAFQKAKERLEAKH 402
DB 362 AMI-----PPTPLPT-----NDVDVYFETSADDNEHAFQKAKERLEAKH 401
QY 403 RERMSQVREWEAEARQAKNLPKADKAVIQHFQKVESLBOEAEANERQQLVETHMARVE 462
DB 402 RNRMRVKKWEAEARQAKNLPKADKAVIQHFQKVESLBOEAEANERQQLVETHMARVE 461
QY 463 AMLNDRRLALENYITALQAVPPRPRHVFNLKKYVRAEQKDROHTLKHFEHVRMVDPK 522
DB 462 AMLNDRRLALENYITALQAVPPRPRHVFNLKKYVRAEQKDROHTLKHFEHVRMVDPK 521
QY 523 AAQTRSQVMTLRLVYIERMNSLLLYNVPVAEEIQDEVDLQKQNTYSDVLANWIS 582
DB 522 AAQTRSQVMTLRLVYIERMNSLLLYNVPVAEEIQDEVDLQKQNTYSDVLANWIS 572
QY 583 EPRISYGNALMPSLTETKTTVELLPVNGEFLSDLOPWHSGADSPANTENEVEPVD 642
DB 573 -----DQFTSSISENPVDVRVSSSESE--EIPPFPLHPF-----PULSENE----- 612


```

Qy 643 RPAADRLGTLTPSGGLTNIKTEI--SEVKMDAEFRHDSGYEVHHQKLVFAEDVGS---- 697
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 613 -----GSGMAEQDG-GLTGABEKVINSKNKMDENNVDETLDV--KEMIFNAERVGLEEE 665
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 698 -----NKGATIGLVGGVVVIATVIVITVLMLKKQYTSIHGGVVVEVDAAVTP 744
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 666 PESVGPUREDFSUSSNALIGLLVIAVAITVIVISLMLKKROYGTSHGIVEVDPMLTP 725
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 745 EERHLKMKQONGENPTYKFFEQM 769
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 726 EERHLKMKQNHGENTPYKLEQM 750
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
A46362
amyloid precursor-like protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C:Accession: A46362
R:Masco, W.; Bupp, K.; Magendantz, M.; Gusella, J.F.; Tanzi, R.E.; Solomon,
Proc. Natl. Acad. Sci. U.S.A. 89, 10758-10762, 1992
A:Title: Identification of a mouse brain cDNA that encodes a protein related
A:Reference number: A46362; MUID:93066322; PMID:1279693
A:Accession: A46362
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-653 <WAS>
A:Experimental source: brain
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:U118683, NCBI:P118684)
C:Superfamily: Alzheimer's disease amyloid beta protein; animal kunitz-type
C:Keywords: transmembrane protein

```

[illegible]

```

Qy    538 YERNQSLSLLYNVPAVAEIQDEVDLQKEONYSDVLANNISSEPRISYGNDAIMP-S 596
      ||||| | | : |: : : ||| : : : : | | : | |
Db    460 EERNQSLGLLDQNPHLAQELRPQIQELL-----LAEHLGPSEL----DASVPGS 505

Qy    597 LTERKTTVELLPNGYESLDDLPWHFSFGADSVPANTENEVEVPDARPAADRGLTRPGS 656
      : | | : : : : | | : : : : | | : : : : |
Db    506 SSEDK-----GSLQP-----PESKDDPPVTLP---KGSTQESS 536

Qy    657 GLTNIKTEETISEVKMDAERHRDSGYEVHH---OKLVFAEDVGNSKGAIIIGMGGVVIA 713
      | : : : : : | : : : : : | : : : : : | : : : : :
Db    537 SSGREKLTPLEQYEQKVNASAPRGFPFHSDDIQDELAPSGTGVSRALSGLLIMAGCG 596

Qy    714 TVIVITLVM-L-KKKQYTSIHGVVVEDVAAYTPPEERHLSKMQQNGENPTYKFEEQ 767
      : |: |: | | | : | | | | | : | | : : : | | | | | : |
Db    597 SLIVLSLLLRRKKPKPYCTISHGVVVEDPMLTLEEQQURELQRHYENPTYRLEE 651

RESULT 11
JC1404
CDEI-box DNA-binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
C:Accession: JC1404
R:Vidal, F.; Blangy, A.; Rassoulzadegan, M.; Cuzin, F.
Biochem. Biophys. Res. Commun. 189, 1336-1341, 1992
A:Title: A murine sequence-specific DNA binding protein shows extensive local similarity
A:Reference number: JC1404; PMID:93129193; PMID:1482349
A:Accession: JC1404
A:Molecule type: mRNA
A:Residues: 1-511 <VID>
C:Comment: This protein plays an important role in the early development of the mouse
C:Keywords: DNA binding; transmembrane protein
```

[illegible]

QY 238 EE-EAADDDEDDGDEVEEAEPEYEEA-TERTTSIAFTTTTTTSEVVEVVEVCSQA 295
Db 297 DSCEGDNTEEDGAGESAEAEVSWDGGAKVVSLSKSSSPSAPAPAPKAPKS 356
QY 296 ETGPCRAMISRMYFDVTEGKCAPFFYGGCGGNRNFDEEYCMAYCGSAMSQSLKTTQE 355
Db 357 ESVTSTPQLS-----ASAAFAVANGSGST-----GAGAPSTAOPTSD 396
QY 356 PLARPVKLPTTAASTPDVADKYETPGDENEHAFQKAKERLEAKHREMSQVWREWE 415
Db 397 P-----YTFHFDHYEQSYKVSQKRLEESHKREKVTVMKWDSD 435
QY 416 AERQAKLPKADKA-----VIOHFQKVESLEQEAANERQOLVETHMARVEAMLNDR 468
Db 436 LEEKYQDMRLADPKAQSFKQRTARFQSVQALEEGNAEKHQLAAMHQQRVLAHINQR 495
QY 469 RLALENYITALQAVPRPRHFMMLKVVRAEQDKRQHTLKHFEH-VRMVDP-----KRAA 524
Db 496 KREAMTCYTOALTEOPPNNAHVHEKCLQKLLRALHDKRAHALAHYRHLNLSGGPGGLEAAA 555
QY 525 QIRSQVMTLHVIRYERMQSLSLYVPAVEI-----QDEV-----562
Db 556 SERPRTLERLIDIRAVNOSMTMLKYPPELSAKIAQLMNDYILALRSKDDIPGSSLGWSE 615
QY 563 -----DELLOKEQNYSDDLVLAN 579
Db 616 EAAGILDKYRVEIERKVAEKELRLAEKQKQRAAREKLEKRLLEAKKVDMLKS 675
QY 580 MISE-----PRISYGNDAIM-----PSUTETKTTVELLPVNG 611
Db 676 QVAEQSQOPTOSTSQAOQOQOQESLPGKELGPDALVTAANPNLETTKS-----726
QY 612 EFLSDDLQPHSGGADSVANTENEVEPDARPAADRGLTTRPGSGLTNKTETSEVKM 671
Db 727 EKLSLDE-----YGEATVSTTKVTQVLPTVDDDAVARVEDVAAA-----VAHQEA 773
QY 672 DAERFDGSGVEVHQKLVF-----FAEDVGSNK-----GAIIGLMVGGVVIATVIVLML 723
Db 774 EPQVQHWTHDLGHRSSFSRLREFQAHAHAKEGRNVVFTLSFACIALMAAVFVGVA 833
QY 724 KKQYTSIH-HGVVEVDAVTP-----ERHLKMQQNGYENPTYKFFPE 766
Db 834 KWRTRSPPHAQGTIEVDQNTVTHPIVREEKIVPNMQINGYENPTYKYFE 883
RESULT 14
S38344
CDEI-binding protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 03-May-1996
C:Accession: S38344
R:Hanes, J.; von der Kammer, H.; Kristiansson, G.I.; Scheit, K.H.
Biochim. Biophys. Acta 1216, 154-156, 1993
A:Title: The complete cDNA coding sequence for the mouse CDEI binding protein.
A:Reference number: S38344; MUID:94032480; PMID:8218408
A:Accession: S38344
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <HAN>
A:Cross-references: EMBL:222592
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase 1

Query Match 17.4%; Score 706; DB 2; Length 246;
Best Local Similarity 51.5%; Pred. No. 5e-31;
Matches 136; Conservative 35; Mismatches 51; Indels 42; Gaps 7;
QY 5 LALLLLAATARALEV-----PTDGNAG-----LLAEPQIAMFCCGRLNMHMVQKWDSDP 56
Db 15 LVLVLLGLTAPAAALAGYIEALAAAGTGFVAEAPQIAMLCGLKLNHNVNIQTGKWEPPD 74
QY 57 SGTCTCIDTKREGILQYQCEVYPELOITNVVEANQPVITQNWCKRGKQCKTHPHVIPYR 116
Db 75 TGTCTGLGTEEVLYQCEYQIEPELQITNVMEANQPVNIDSWCRDRKQCKS--HIVIPFK 132

QY 117 CLVGEFYSALLVPDKCKFLHOERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGI 176
Db 133 CLVGEFYSVLLVPDNCQFFHQRMEYCEKQRHWHITLVKEACLTGLTLYSGMLLPCGV 192
QY 177 DKFRGVFVCCPLAEESDNVSDADEDDSDVWVGADTDYADGSDGKVVVEVAEEBVAE 236
Db 193 DQFHGTGVCCP---QTKTVDS-----DSTMSKEEEE---222
QY 237 VEDEEADDED-DEGDGEVEEEAE 259
Db 223 -EEDEEEDDYDLDKSEFPTEAD 245
RESULT 15
A32282
Alzheimer's disease amyloid beta protein precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Aug-1989 #sequence_revision 17-Aug-1989 #text_change 13-Aug-1999
C:Accession: A32282
R:Yamada, T.; Sakaki, H.; Dohura, K.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 158, 906-912, 1989
A:Title: Structure and expression of the alternatively-spliced forms of mRNA for the
A:Reference number: A32282; MUID:89149813; PMID:2493250
A:Accession: A32282
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-100 <YAM>
A:Cross-references: GB:M24397; NID:g200350; PIDN:AAA39929, 1; PID:g200351
C:Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase
C:Keywords: alternative splicing
F:11-61/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
Query Match 12.3%; Score 501; DB 2; Length 100;
Best Local Similarity 92.0%; Pred. No. 1.5e-20;
Matches 92; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 281 ESVEEVVREVCSEQAETGPCRAMISRMYFDVTEGKCAPFFYGGCGGNRNFDTBEYCMAY 340
Db 1 ESVEEVVREVCSEQAETGPCRAMISRMYFDVTEGKCVPFYGGCGGNRNFDTBEYCMAY 60
QY 341 CGSAMSQSLKTTQEPPLARDPVKLPPTTAASTPDVADVKYLE 380
Db 61 CGSVSTQSLKTTSEPLPDQPKLPTTAASTPDVADVKYLE 100
Search completed: January 28, 2003, 14:19:59
Job time : 25.1912 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 11.7782 seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLAATRALEV.....KMQQNGYENPTYKFFEQMN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	1 A4_HUMAN	P05067 homo sapien
2	3937	97.0	770	1 A4_RAT	P08592 rattus norv
3	3927	96.8	770	1 A4_MOUSE	P12023 mus musculu
4	3921.5	96.6	751	1 A4_SAIISC	Q95241 saimiri sci
5	1981.5	48.8	763	1 APP2_HUMAN	Q06481 homo sapien
6	1971.5	48.6	765	1 APP2_RAT	P15943 rattus norv
7	1704.5	42.0	695	1 APP2_MOUSE	Q06335 mus musculu
8	1155.5	28.5	650	1 APPI_HUMAN	P11693 homo sapien
9	1150.5	28.4	653	1 APPI_MOUSE	Q03157 mus musculu
10	785	19.3	686	1 A4_CAEEL	Q10651 caenorhabdi
11	737.5	18.2	886	1 A4_DROME	P14599 drosophila
12	457	11.3	87	1 A4_MACFA	P53601 macaca fasc
13	403	9.9	76	1 A4_MACMU	P29216 macaca mula
14	292	7.2	59	1 A4_BOVIN	Q28053 bos taurus
15	288	7.1	58	1 A4_RABIT	Q28748 oryctolagus
16	288	7.1	58	1 A4_SHEEP	Q28757 ovis aries
17	287	7.1	58	1 A4_CANFA	Q28280 canis famil
18	283	7.0	57	1 A4_PIG	Q29023 sus scrofa
19	283	7.0	57	1 A4_URUMA	Q29149 ursus marit
20	194.5	4.8	3911	1 AKA9_HUMAN	Q39996 h a-kinase
21	186	4.6	252	1 SPT2_HUMAN	O43278 homo sapien
22	186	4.6	993	1 SCPI_MOUSE	Q62209 mus musculu
23	176	4.3	55	1 ISH1_STOHE	P31713 stoichactis
24	175.5	4.3	579	1 G160_HUMAN	Q08378 homo sapien
25	175	4.3	302	1 TPPI_RAT	Q02445 rattus norv
26	174.5	4.3	252	1 SPT2_MOUSE	Q3wu03 mus musculu
27	174.5	4.3	513	1 SPT1_HUMAN	O43278 homo sapien
28	169	4.2	1875	1 MLF1_YEAST	Q02455 saccharomyc
29	168	4.1	304	1 TPPI_HUMAN	P10646 homo sapien
30	166	4.1	55	1 ISH2_STOHE	P81129 stoichactis
31	164.5	4.1	507	1 SPT1_MOUSE	Q39097 mus musculu
32	163.5	4.0	346	1 ANBP_MERUN	Q25577 meriones un
33	163.5	4.0	1130	1 YL17_CAEEL	Q11102 caenorhabdi

RESULT 1

ID	A4_HUMAN	STANDARD;	PRT;	770 AA.
AC	P05067; P09000; Q16011;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alzheimer's disease amyloid A4 protein precursor (Protease nexin-II)			
DE	(PN-II) (APPI) (Contains: Beta-amyloid protein (Beta-APP) (A-beta))			
GN	APP OR A4 OR CVAP OR ADL			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=671144572; PubMed=2881207;			
RA	Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,			
RA	Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;			
RT	"The precursor of Alzheimer's disease amyloid A4 protein resembles a			
RT	cell-surface receptor.";			
RL	Nature 325:733-736(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88122639; PubMed=2893289;			
RA	Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,			
RA	Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,			
RA	Cordell B.;			
RT	"A new A4 amyloid mRNA contains a domain homologous to serine			
RT	proteinase inhibitors.";			
RL	Nature 331:525-527(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89128427; PubMed=2783775;			
RA	Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,			
RA	Unterbeck A., Beyreuther K., Mueller-Hill B.;			
RT	"The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid			
RT	is encoded by 16 exons.";			
RL	Nucleic Acids Res. 17:517-522(1989).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97263807; PubMed=9108164;			
RA	Hattori M., Tsukahara F., Furuhashi Y., Tanahashi H., Hirose M.,			
RA	Saito M., Tsukuni S., Sakaki Y.;			
RT	"A novel method for making nested deletions and its application for			
RT	sequencing of a 300 kb region of human APP locus.";			
RL	Nucleic Acids Res. 25:1802-1808(1997).			
RN	[5]			
RP	SEQUENCE OF 286-345 AND 365-366 FROM N.A.			
RX	MEDLINE=88122640; PubMed=2893290;			
RA	Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,			
RA	Gusella J.F., Neve R.L.;			
RT	"Protease inhibitor domain encoded by an amyloid protein precursor			
RT	mRNA associated with Alzheimer's disease.";			
RL	Nature 331:528-530(1988).			

ALIGNMENTS

34	163	4.0	2944	1	CA17_HUMAN	Q02388 homo sapien
35	163	4.0	3176	1	CA36_HUMAN	P12111 homo sapien
36	162.5	4.0	297	1	TRT2_HUMAN	P45379 homo sapien
37	162	4.0	304	1	TFPI_MACMU	Q28864 macaca mula
38	162	4.0	407	1	IE68_HSVSA	Q01042 herpesvirus
39	162	4.0	3137	1	CA36_CHICK	P15989 gallus gall
40	161.5	4.0	630	1	YCF2_OENVI	P31569 oenothera v
41	161.5	4.0	1278	1	DYNA_HUMAN	Q14203 homo sapien
42	161	4.0	64	1	SPT3_HUMAN	P49223 homo sapien
43	160.5	4.0	721	1	YCF2_OENPI	P31568 oenothera p
44	160	3.9	802	1	NAB3_YEAST	P38996 saccharomyc
45	159	3.9	197	1	MCPI_MELCP	P92968 melichaea c

RN [16] SEQUENCE OF 287-367 FROM N.A.
 RP MEDLINE=88122641; PubMed=2893291;
 RX Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
 RA "Novel precursor of Alzheimer's disease amyloid protein shows
 RT protease inhibitory activity.";
 RL Nature 331:530-532(1988).
 RN [17]
 RP SEQUENCE OF 284-289 AND 365-770 FROM N.A.
 RX MEDLINE=87231971; PubMed=3035574;
 RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
 RT "Molecular cloning and characterization of a cDNA encoding the
 RL cerebrovascular and the neuritic plaque amyloid peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
 RN [18]
 RP SEQUENCE OF 507-770 FROM N.A.
 RX MEDLINE=88124954; PubMed=2893379;
 RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
 RA Marotta C.A.;
 RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
 RL disease brain: coding and noncoding regions of the fetal precursor
 RT mRNA are expressed in the cortex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
 RN [19]
 RP SEQUENCE OF 672-681.
 RX MEDLINE=88035004; PubMed=3312495;
 RA Partridge W.M., Vinters H.V., Yang J., Eisenberg J., Choi T.B.,
 RA Tourtellotte W.W., Huebner V., Shively J.E.;
 RT "Amyloid angiopathy of Alzheimer's disease: amino acid composition
 RL and partial sequence of a 4,200-dalton peptide isolated from cortical
 RT microvessels.";
 RL J. Neurochem. 49:1394-1401(1987).
 RN [10]
 RP SEQUENCE OF 739-770 FROM N.A.
 RX MEDLINE=90236318; PubMed=2110105;
 RA Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
 RT "Genomic organization of the human amyloid beta-protein precursor
 RL gene.";
 RL Gene 87:257-263(1990).
 RN [11]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=89016647; PubMed=3140222;
 RA Schon E.A., Mita S., Sadlock J., Herbert J.;
 RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
 RL encodes a 95-kDa polypeptide.";
 RL Nucleic Acids Res. 16:9351-9351(1988).
 RN [12]
 RP SEQUENCE OF 18-50.
 RX MEDLINE=87250462; PubMed=3597385;
 RA van Nostrand W.E., Cunningham D.D.;
 RT "Purification of protease nexin II from human fibroblasts.";
 RL J. Biol. Chem. 262:8508-8514(1987).
 RN [13]
 RP IDENTITY OF APP WITH NEXIN-II.
 RX MEDLINE=89384866; PubMed=2506449;
 RA Oltersdorf T., Fritz L.C., Schenk D.B., Lieberburg I.,
 RA Johnson-Wood K.L., Beattie E.C., Ward P.J., Blacher R.W., Dovey H.F.,
 RA Sinha S.;
 RT "The secreted form of the Alzheimer's amyloid precursor protein with
 RL the Kunitz domain is protease nexin-II.";
 RL Nature 341:144-147(1989).
 RN [14]
 RP PROTEASE-SPECIFICITY OF INHIBITOR DOMAIN.
 RX MEDLINE=90211252; PubMed=1969731;
 RA Kido H., Fukutomi A., Schilling J., Wang Y., Cordell B., Katunuma N.;
 RT "Protease-specificity of Kunitz inhibitor domain of Alzheimer's
 RL disease amyloid protein precursor.";
 RL Biochem. Biophys. Res. Commun. 167:716-721(1990).
 RN [15]
 RP COMPLEX WITH G(O).
 RX MEDLINE=93188965; PubMed=8446172;
 RA Nishimoto-I., Okamoto T., Matsuura Y., Takahashi S., Okamoto T.,

RA Murayama Y., Ogata E.;
 RT "Alzheimer amyloid protein precursor complexes with brain GTP-binding
 RL protein G(O).";
 RL Nature 362:75-79(1993).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 28-133.
 RX MEDLINE=99215582; PubMed=10201399;
 RA Rossjohn J., Cappai R., Feil S.C., Henry A., McKinstry W.J.,
 RA Galatis D., Hesse L., Multhaup G., Beyreuther K., Masters C.L.,
 RA Parker M.W.;
 RT "Crystal structure of the N-terminal, growth factor-like domain of
 RL Alzheimer amyloid precursor protein.";
 RL Nat. Struct. Biol. 6:327-331(1999).
 RN [17]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 287-344.
 RX MEDLINE=91104913; PubMed=2125487;
 RA Hynes T.R., Randal M., Kennedy L.A., Eigenbrot C., Kossiakof A.A.;
 RT "X-ray crystal structure of the protease inhibitor domain of
 RL Alzheimer's amyloid beta-protein precursor.";
 RL Biochemistry 29:10018-10022(1990).
 RN [18]
 RP STRUCTURE BY NMR OF 289-344.
 RX MEDLINE=92031488; PubMed=1718421;
 RA Heald S.L., Tilton R.F. Jr., Hammond L.S., Lee A., Bayney R.M., A.,
 RA Kamark M.E., Ramabhadran T.V., Dreyer R.N., Davis G., Unterbeck A.,
 RA Tamburini P.P.;
 RT "Sequential NMR resonance assignment and structure determination of
 RL the Kunitz-type inhibitor domain of the Alzheimer's beta-amyloid
 RL precursor protein.";
 RL Biochemistry 30:10467-10478(1991).
 RN [19]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=94281210; PubMed=7516706;
 RA Talafous J., Marciniowski K.J., Klopman G., Zagorski M.G.;
 RT "Solution structure of residues 1-28 of the amyloid beta-peptide.";
 RL Biochemistry 33:7788-7796(1994).
 RN [20]
 RP STRUCTURE BY NMR OF 696-706.
 RX MEDLINE=97128622; PubMed=8973180;
 RA Kohno T., Kobayashi K., Maeda T., Sato K., Takashima A.;
 RT "Three-dimensional structures of the amyloid beta peptide (25-35) in
 RL membrane-mimicking environments.";
 RL Biochemistry 35:16094-16104(1996).
 RN [21]
 RP STRUCTURE BY NMR OF 672-711.
 RX MEDLINE=98359783; PubMed=9693002;
 RA Coles M., Bicknell W., Watson A.A., Fairlie D.P., Craik D.J.;
 RT "Solution structure of amyloid beta-peptide(1-40) in a water-micelle
 RL environment. Is the membrane-spanning domain where we think it is?";
 RL Biochemistry 37:11064-11077(1998).
 RN [22]
 RP STRUCTURE BY NMR OF 672-699.
 RX MEDLINE=20400066; PubMed=10940222;
 RA Poulsen S.-A., Watson A.A., Craik D.J.;
 RT "Solution structures in aqueous SDS micelles of two amyloid beta
 RL peptides of Abeta(1-28) mutated at the alpha-secretase cleavage
 RL site.";
 RL J. Struct. Biol. 130:142-152(2000).
 RN [23]
 RP STRUCTURE BY NMR OF 681-706.
 RX MEDLINE=20400065; PubMed=10940221;
 RA Zhang S., Iwata K., Lachenmann M.J., Peng J.W., Li S., Stimson E.R.,
 RA Lu Y., Felix A.M., Maggio J.E., Lee J.P.;
 RT "The Alzheimer's peptide a beta adopts a collapsed coil structure in
 RL water.";
 RL J. Struct. Biol. 130:130-141(2000).
 RN [24]
 RP SIGNAL SEQUENCE CLEAVAGE SITE, AND TOPOLOGY.
 RX MEDLINE=88296437; PubMed=2900137;
 RA Dyrks T., Weidemann A., Multhaup G., Salbaum J.M., Lemaire H.-G.,
 RA Kang J., Mueller-Hill B., Masters C.L., Beyreuther K.;
 RT "Identification, transmembrane orientation and biogenesis of the
 RL amyloid A4 precursor of Alzheimer's disease."

Query Match		100.0%;	Score 4058;	DB 1;	Length 770;
Best Local Similarity		100.0%;	Pred. No. 3.1e-204;		
Matches	770;	Conservative	0;	Mismatches	0;
				Indels	Gaps
QY	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNHHMVQNGKWDSPSGTK	60		
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNHHMVQNGKWDSPSGTK	60		
QY	61	TCIDTREGILQYCOEYVPELOITNVVEANQPVTIQNCCKRGKCKOCTHPHFVTPYRCLVG	120		
Db	61	TCIDTREGILQYCOEYVPELOITNVVEANQPVTIQNCCKRGKCKOCTHPHFVTPYRCLVG	120		
QY	121	EFVSDALLVPDKCFKFLHQERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKPR	180		
Db	121	EFVSDALLVPDKCFKFLHQERMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKPR	180		
QY	181	GVFVCCPLAESDNVDSADAEDDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE	240		
Db	181	GVFVCCPLAESDNVDSADAEDDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEEVEE	240		
QY	241	EADDDDEDDGDEVEEAEPEYEATERTSTATTTTTTSTESVEEVVVCSEQAETGPC	300		
Db	241	EADDDDEDDGDEVEEAEPEYEATERTSTATTTTTTSTESVEEVVVCSEQAETGPC	300		
QY	301	RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD	360		
Db	301	RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD	360		
QY	361	PVKLPPTAASDPADVKYLETGPDENEHAHFQAKERLEAKHIREMSQVMEWEAEARQA	420		
Db	361	PVKLPPTAASDPADVKYLETGPDENEHAHFQAKERLEAKHIREMSQVMEWEAEARQA	420		
QY	421	KNLPKADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480		
Db	421	KNLPKADKKAIVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL	480		
QY	481	QAVPPRRHVNMLKKYVRAEQKDRQHTLKFHEHVRMVDPKKAAQIRSQVMTHLRVIYER	540		
Db	481	QAVPPRRHVNMLKKYVRAEQKDRQHTLKFHEHVRMVDPKKAAQIRSQVMTHLRVIYER	540		
QY	541	MNQSLLYNVPAVEIQDEVDDELQKQNSYDDVLANNISEPRISYNDALMPSLTET	600		
Db	541	MNQSLLYNVPAVEIQDEVDDELQKQNSYDDVLANNISEPRISYNDALMPSLTET	600		
QY	601	KTVVELLPVNGEFLDLPWHSFGADSVPAENTENEVEPVDARPAADRLGTLTRPGSLTN	660		
Db	601	KTVVELLPVNGEFLDLPWHSFGADSVPAENTENEVEPVDARPAADRLGTLTRPGSLTN	660		
QY	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVIATVIVITL	720		
Db	661	IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVIATVIVITL	720		
QY	721	VMLKKKQYTSIHGVEVDAVAAVTPERHLSKMQQNGYENPTYKFFEQMQN	770		
Db	721	VMLKKKQYTSIHGVEVDAVAAVTPERHLSKMQQNGYENPTYKFFEQMQN	770		
RESULT 2					
A4_RAT	A4_RAT	STANDARD;	PRT:	770 AA.	
AC	P08592;				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Alzheimer's disease amyloid A4 protein homolog precursor				
DE	(Amyloidogenic glycoprotein) (AG).				
GN	APP.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_Taxid=10116;				
RN	[1]				

SEQUENCE OF 1-289 AND 365-770 FROM N.A.
TISSUE=Brain;
RX MEDLINE=88312583; PubMed=2900758;
RA Shivers B.D., Hilbich C., Mulchaup G., Salbaum J.M., Beyreuther K.,
Seeburg P.H.;
RT "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
in rat brain suggests a role in cell contact.";
RL EMBO J. 7:1365-1370(1988).
RN [2]
RP SEQUENCE OF 289-364 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89183625; PubMed=2648331;
RA Kang J., Mueller-Hill B.;
RT "The sequence of the two extra exons in rat preA4.";
RL Nucleic Acids Res. 17:2130-2130(1989).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
SPLICING.
CC -1- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION
WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC
RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE
NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF
PHOSPHORYLATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL; X07648; CAA30488.1; -;
DR EMBL; X14066; CAA32229.1; -;
DR PIR; S00550; S00550.
DR PIR; S03607; S03607.
DR HSSP; P05067; IAAP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta_APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_Bpti; 1.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta_APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR PRODOM; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
KW Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW Alternative splicing; Serine protease inhibitor.
FT SIGNAL 1 17
FT CHAIN 18 770
FT ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT HOMOLOG.
FT DOMAIN 18 699
FT TRANSMEM 700 723
FT DOMAIN 724 770
FT DOMAIN 673 715
FT DOMAIN 287 345
FT SITE 759 762
FT DISULFID 291 341
FT DISULFID 300 324
FT DISULFID 316 337
FT CARBOHYD 542 542
FT CARBOHYD 571 571
FT VARSPLIC 289 289
FT VARSPLIC 290 364
FT MISSING (IN ISOFORM APP(695)).

DR PIR; A27485; A27485.
 DR PIR; S04855; S04855.
 DR PIR; S19727; S19727.
 DR HSP; P05067; IAA.
 DR MGD; MGI:88059; App.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPDASE.
 DR ProDom; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neutone; Transmembrane; Signal;
 KW Alternative splicing; Serine protease inhibitor.
 FT SIGNAL 1 17
 FT CHAIN 18 770
 FT
 FT DOMAIN 18 699
 FT TRANSMEM 700 723
 FT DOMAIN 724 770
 FT DOMAIN 673 715
 FT SITE 759 762
 FT DISULFID 291 341
 FT DISULFID 300 324
 FT DISULFID 316 337
 FT CARBOHYD 542 542
 FT CARBOHYD 571 571
 FT VARSPLIC 289 289
 FT VARSPLIC 290 364
 FT VARSPLIC 346 380
 SQ SEQUENCE 770 AA; 86752 MW; 26C50DE0890CAF7A CRC64;

Query Match 96.8%; Score 3927; DB 1; Length 770;
 Best Local Similarity 96.8%; Pred. No. 2.1e-197;
 Matches 745; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIANFCGLNHHMVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPQIANFCGLNHHMVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILQYCOEYVPELQINNVYEANQPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 DB 61 TCIDTREGILQYCOEYVPELQINNVYEANQPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 QY 121 EFVSDALLVPDKCFELHQRMDYCYETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCFELHQRMDYCYETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GFVEVCCPLAEDSDVSDAEDDDSDVWVGADTDYAGSEDKVYVEVAEEVEE 240
 DB 181 GFVEVCCPLAEDSDVSDAEDDDSDVWVGADTDYAGSEDKVYVEVAEEVEE 240
 QY 241 EADDEDEDGDEVEEAEPEYEATRTTSIATTTTTTTEVEEYVREVCSEQAGTGPC 300
 DB 241 EADDEDEDGDEVEEAEPEYEATRTTSIATTTTTTTEVEEYVREVCSEQAGTGPC 300
 QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQPELARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQPELARD 360
 QY 361 PVKLPTTAATPDPAVKYLETGDENEHAHFQAKERLEAKHRMSQVNRWEAEERQA 420
 DB 361 PVKLPTTAATPDPAVKYLETGDENEHAHFQAKERLEAKHRMSQVNRWEAEERQA 420

QY 421 KNLPRADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
 DB 421 KNLPRADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480
 QY 481 QAVPPRPFRHVNMLKKYVRAEQKDRQHTLKHFHEHVMYVDPKKAQIRSQVTHLRIYIER 540
 DB 481 QAVPPRPFRHVNMLKKYVRAEQKDRQHTLKHFHEHVMYVDPKKAQIRSQVTHLRIYIER 540
 QY 541 MNQSLSLLYNYPVAAEETQDEVDLLOKEQYSDVLANMISEPRISYGNLALMPSLTET 600
 DB 541 MNQSLSLLYNYPVAAEETQDEVDLLOKEQYSDVLANMISEPRISYGNLALMPSLTET 600
 QY 601 KTTVELLPVNGEFSLDDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTRPCSGLTN 660
 DB 601 KTTVELLPVNGEFSLDDLPWHPFGVDSPVPAANTENEVEPVDARPAADRGLTRPCSGLTN 660
 QY 661 IKTEEISEVKMDAEPHRDGYEVHQQKLVFFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 DB 661 IKTEEISEVKMDAEPHRDGYEVHQQKLVFFAEDVGSNGKGAIIGLMVGGVVIATVITL 720
 QY 721 VMLKKQYTSIHGVEVVDAAVTPPEERHLSKMQNGYENPTYKFEQMN 770
 DB 721 VMLKKQYTSIHGVEVVDAAVTPPEERHLSKMQNGYENPTYKFEQMN 770

RESULT 4
 A4_SAISC
 ID A4_SAISC STANDARD; PRT; 751 AA.
 AC Q95241;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein precursor [Contains: Beta-amyloid protein (Beta-APP) (A-beta)].
 GN APP.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
 OX NCBI_TaxID=9521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Kidney;
 RX MEDLINE=96108492; PubMed=8532114;
 RA Levy E., Amorim A., Frangione B., Walker L.C.;
 RT "Beta-amyloid precursor protein gene in squirrel monkeys with cerebral amyloid angiopathy";
 RL Neurobiol. Aging 16:805-808(1995).
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN G(O).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DOMAIN: THE CLATHRIN-BINDING SITE IS ESSENTIAL FOR ITS ASSOCIATION WITH X11-ALPHA, -BETA, AND -GAMMA. THE SEQUENCE SPECIFIC RECOGNITION EXTENDS TO PEPTIDE RESIDUES THAT ARE C-TERMINAL TO THE NPXY MOTIF. THIS INTERACTION APPEARS TO BE INDEPENDENT OF PHOSPHORYLATION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S81024; AAD14347.1; -.
 DR HSP; P05067; IAA.
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR Pfam; PF00014; Kunitz_BPTI; 1.

DR Pfam: PF02177; A4_EXTRA; 1.
 DR DR PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOIDA4.
 DR PRINTS; PR00759; BASICPTASE.
 DR PRODOM; PD000222; Kunitz_BPTI; 1.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR SMART; SM00131; KU; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Transmembrane; Alternative splicing;
 KW Signal; Serine protease inhibitor.
 FT SIGNAL 1 17 BY SIMILARITY.
 FT CHAIN 18 751 A4 PROTEIN.
 FT CHAIN 653 695 BETA-AMYLOID PROTEIN (POTENTIAL).
 FT DOMAIN 18 680 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 681 704 POTENTIAL.
 FT DOMAIN 705 751 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 287 345 BPTI/KUNITZ INHIBITOR.
 FT SITE 740 743 CLATHRIN-BINDING (BY SIMILARITY).
 FT ACT_SITE 301 302 REACTIVE BOND.
 FT DISULFID 291 341 BY SIMILARITY.
 FT DISULFID 300 324 BY SIMILARITY.
 FT DISULFID 316 337 BY SIMILARITY.
 FT CARBOHYD 523 523 N-LINKED (GLCNAC. .) (PROBABLE).
 FT CARBOHYD 552 552 N-LINKED (GLCNAC. .) (PROBABLE).
 SQ SEQUENCE 751 AA; 84893 MW; 6C3E431089569049 CRC64;

Query Match 96.6%; Score 3921.5; DB 1; Length 751;
 Best Local Similarity 96.9%; Pred. No. 3.8e-197;
 Matches 746; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 1 MFLGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGLRNHNMVQNGKWDSPSGTK 60
 DB 1 MFLGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGLRNHNMVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILOVCOEYPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVYRCVLG 120
 DB 61 TCIDTREGILOVCOEYPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVYRCVLG 120
 QY 121 EFVSDALLVPDKCKFLHQRMVDCVTHLHWHVAKETCEKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMVDCVTHLHWHVAKETCEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAEEDSDNVDADADDSDVWVGADTDYADGSKVVEVAEEVEEVEE 240
 DB 181 GVEFVCCPLAEEDSDNVDADADDSDVWVGADTDYADGSKVVEVAEEVEEVEEVEE 240
 QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSEVEEYVREVCSEQAETGPC 300
 DB 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSEVEEYVREVCSEQAETGPC 300
 QY 301 RAMISRWYFDVTGKCAPFFYCGCGGNRRNFDTEECVMAVCSSMSQSLLKTKTQEP LARD 360
 DB 301 RAMISRWYFDVTGKCAPFFYCGCGGNRRNFDTEECVMAVCSSMSQSLLKTKTQEP LARD 360
 QY 361 PVKLPTAASTPDVADVKYLETGDENEHAHFQKAKERLEAKHRMSQVMRWEAEAEQA 420
 DB 361 PVKLPTAASTPDVADVKYLETGDENEHAHFQKAKERLEAKHRMSQVMRWEAEAEQA 420
 QY 421 KNLPKADKAVIQHFOGKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 DB 421 KNLPKADKAVIQHFOGKVESLEQEAANERQQLVETHMARVEAMLNDRRRRLALENYITAL 480
 QY 481 QAVPPRRHVNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKKAQIRSQVTHLRYVIER 540
 DB 481 QAVPPRRHVNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKKAQIRSQVTHLRYVIER 540
 QY 541 MNGSLSLYNVPAVAEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
 DB 541 MNGSLSLYNVPAVAEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
 QY 552 MNGSLSLYNVPAVAEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTLET 581
 DB 552 MNGSLSLYNVPAVAEIQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTLET 581

QY 601 KTTVELLPVNGEFSLLDLPWHSGADSVDPANTENEVPDARPAADRLTRPGSLTN 660
 DB 582 KTTVELLPVNGEFSLLDLPWHSGADSVDPANTENEVPDARPAADRLTRPGSLTN 641
 QY 661 IKTEIIESEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIGLMVGGVVIAVITL 720
 DB 642 IKTEIIESEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIGLMVGGVVIAVITL 701
 QY 721 VMLKKQYTSIHGVEVDAAVTPERHLSKMQONGYENPTYKFEQMON 770
 DB 702 VMLKKQYTSIHGVEVDAAVTPERHLSKMQONGYENPTYKFEQMON 751

RESULT 5
 APP2_HUMAN STANDARD; PRT; 763 AA.
 ID APP2_HUMAN
 AC Q06481;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid-like protein 2 precursor (Amyloid protein homolog) (APPH)
 DE (CDEI-box binding protein) (CDEBP).
 GN APLP2 OR APLP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA MEDLINE=93250009; PubMed=8485127;
 RX Sprecher C.A., Grant F.J., Grimm G., O'Hara P.J., Norris F.,
 RA Norris K., Foster D.C.;
 RT "Molecular cloning of the cDNA for a human amyloid precursor protein
 RT homolog: evidence for a multigene family.";
 RL Biochemistry 32:4481-4486(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA MEDLINE=9521734; PubMed=7702756;
 RX von der Kammer H., Hanes J., Klaudiny J., Scheit K.H.;
 RT "A human amyloid precursor-like protein is highly homologous to a
 RT mouse sequence-specific DNA-binding protein.";
 RL DNA Cell Biol. 13:1137-1143(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=94035131; PubMed=8220435;
 RX Wasco W., Gurubagavatula S., Paradis M., Romano D.M., Sisodia S.S.,
 RA Hyman B.T., Neve R.L., Tanzi R.E.;
 RT "Isolation and characterization of APLP2 encoding a homologue of the
 RT Alzheimer's associated amyloid beta protein precursor.";
 RL Nat. Genet. 5:93-99(1993)
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF HEMOSTASIS. THE
 CC SOLUBLE FORM MAY HAVE INHIBITORY PROPERTIES TOWARDS COAGULATION
 CC FACTORS. MAY INTERACT WITH CELLULAR G-PROTEIN SIGNALING PATHWAYS.
 CC MAY BIND TO THE DNA 5'-GTCACATG-3'(CDEI BOX).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
 CC (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2 AND
 CC 3; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: IN PLACENTA, BRAIN, HEART, LUNG, LIVER, KIDNEY
 CC AND ENDOTHELIAL TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.)


```

CC EMBL; S60099; AAC60589.1; -
DR EMBL; L09209; AAA35526.1; -
DR EMBL; 225572; CAA80295.1; -
DR EMBL; L27631; CAA41701.1; -
DR HSP; P05067; IMWP.
DR Genew; HGNC:598; APLP2.
DR MIM; 104776; -
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PS00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI. 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Transmembrane; Signal; Alternative splicing; DNA-binding;
KW Nuclear protein; Serine protease inhibitor.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 763 AMYLOID-LIKE PROTEIN 2.
FT DOMAIN 30 692 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 693 716 POTENTIAL.
FT DOMAIN 717 763 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 215 280 ASP/GLU-RICH (HIGHLY ACIDIC).
FT DOMAIN 306 364 BPTI/KUNITZ INHIBITOR.
FT DOMAIN 215 231 POLY-GLU.
FT ACT_SITE 320 321 REACTIVE BOND (BY SIMILARITY).
FT DISULFID 310 360 BY SIMILARITY.
FT DISULFID 319 343 BY SIMILARITY.
FT DISULFID 335 356 BY SIMILARITY.
FT VARSPPLIC 308 363 MISSING (IN ISOFORM 2).
FT VARSPPLIC 613 624 MISSING (IN ISOFORM 3).
FT CONFLICT 543 543 S -> I (IN REF. 1).
SQ SEQUENCE 763 AA; 86955 MW; CA3A7D6DDB8A28D0 CRC64;

Query Match 48.8%; Score 1981.5; DB 1; Length 763;
Best Local Similarity 50.4%; Pred. No. 3.6e-96;
Matches 408; Conservative 127; Mismatches 168; Indels 107; Gaps 21;

QY 5 LALLLLAAWTAARALEY-----PTDGNAG---LLAEPQIAWFCGRLLNMHNMVONGKWDSDP 56
DB 15 LLLLLLGLTAPALAGYIEALANAGTGFVAEPQIAWFCGRLLNMHNMVONGKWDSDP 74
QY 57 SGTKTCTDKTGELQYCEYVPELOITNVVEANQPTVIONWCKRGKCKTTPHFEVPIYR 116
DB 75 TGTKSCFETKEEVLYQCQEMYPELOITNVVEANQVRSIDNWCRRDKCKS--RFVTPPK 132
QY 117 CLVGEFVSALLVDPKCKFLQERNDVCETHLHWHVAVKETSCKTNLHDYGMLLPCGI 176
DB 133 CLVGEFVSALLVDPKCKFLQERNDVCETHLHWHVAVKETSCKTNLHDYGMLLPCGI 192
QY 177 DKRGVEFVCCPLAESDNVDSADADDSDVWVGADTDYADGSDKDVVEVAEEVEAE 236
DB 193 DQFHGTGYVCCPOTKIGSVSKREEDEE-----EEEEDEEDYVYKSEFPTTEAD 245
QY 237 VEE--EEA--DDDEDDGDEVEEAEE-----EPYEATERITTSIATTTTTSYVE 284
DB 246 LEDTEAAVDEDDDEDEEVEEDRDYVDFKGDYNE--ENPTEPGSDGTMDSKEIT 303
QY 285 EVREYVCSQAEAGPCRAMISRHYFDVTGKCAPFFYGGCGGNRNNFDTVEECVACGSA 344
DB 304 HDVKAVCSQAEAMTGPCRAMPRYFDLSKGVRFYIGCGGNRNNFDESDYCMVCKAM 363
QY 345 MSQSLLKTTQPLAROPVKLTPTTAASTPDVADKYLETGDNENHAFQKAKERLEAKHRE 404
DB 364 I-----PPTPLPT-----NDVDFYFTSADNNEHAFQKAKERLEAKHRE 403
QY 405 RNSQVMREWEAEARQAKNLPKADKAVIQHFQEKVESLEQEAEANERQQLVETHMARVEAM 464

```

```

DB 404 RMDRVKKEWEAEALQAKNLPRAERQTLQHFQAMVKALEKAASEKQOLVETHLARVEAM 463
QY 465 LNDRRRLALENYITALQAVPPRPRIHVFNMLKKYVRAEQKDRQHTLKHFHEHVMVDPKKAA 524
DB 464 LNDRRRLALENYIALQALQSDPPRPHRIILQALRRYVRAENKDRLHTIRHYQHVLAVDPEKAA 523
QY 525 QIRSOVMTHLRVIERMNSLSLLYNVPAVAEETODEVDELLQKQENSDVDDVLANMISEP 584
DB 524 QMKSQVMTHLRVIERMNSLSLLYNVPAVAEETODEVDELLQKQENSDVDDVLANMISEP 572
QY 585 RISYGNDAIMPSTETKTVELLPVNGEFSLLDLPQWHSFGADSVDPANTNEVEPVDARP 644
DB 573 -----DQFTASISETPVDVR---VSSEES-EETPPFPF--HPEPALPENE-----DTQP 616
QY 645 AADRLTTRPGSLTN-----IKTEE-----ISEVKMDAEFRHDSGYEVHHQKLVFAEDVG 696
DB 617 ELYHPM--KGGSGVGEQDGLIGAEKVINSKKNVDENMVIDETLDV--KEMIFNAERVG 672
QY 697 S-----NKGAILGLMYGGVVIATVITLVMLKKKQYTSIHHCVVVEVD 739
DB 673 GLEERESVGPLREDFSLSSALIGLLVIAVATVIVISLVMLKKRQYGTISHGIVEVD 732
QY 740 AAVTPEERHLSKMOQNGYENPTYKFFEQMQ 769
DB 733 PMLTPEERHLNKMOMHNGYENPTYKYLEQMQ 762

RESULT 6
APP2_RAT
ID APP2_RAT STANDARD; PRT; 765 AA.
AC P15943;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Amyloid-like protein 2 precursor (Sperm membrane protein YWK-II).
GN APLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RC SEQUENCE OF 1-627 FROM N.A.
RP STRAIN=Wistar; TISSUE=Brain, and Heart;
RX MEDLINE=94336849; PubMed=8086458;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "Complete nucleotide and deduced amino acid sequence of rat amyloid
RT protein precursor-like protein 2 (APLP2/APPH): two amino acids length
RT difference to human and murine homologues.";
RL Biochim. Biophys. Acta 1219:167-170(1994).
RN [2]
RP SEQUENCE OF 575-765 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=90207205; PubMed=1690887;
RA Yan Y.C., Bai Y., Wang L.F., Miao S.Y., Koide S.S.;
RT "Characterization of cDNA encoding a human sperm membrane protein
RT related to A4 amyloid protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2405-2408(1990).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B, C AND D;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DB EMBL; X77934; CAA54906.1; -.

```



```

DR EMBL; M31322; AAA42352.1; -.
DR PIR; A35981; A35981.
DR HSSP; P05067; 1MWP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR Pfam; PF02177; A4_EXTRA; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SK00006; A4_EXTRA; 1.
DR SMART; SK00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS00279; BPTI_KUNITZ_2; 1.
KW Transmembrane; Alternative splicing; Serine protease inhibitor;
KW Signal; Glycoprotein.
FT SIGNAL 1 29
FT CHAIN 30 765
FT DOMAIN 30 695
FT TRANSMEM 696 718
FT DOMAIN 719 765
FT DOMAIN 218 282
FT DOMAIN 308 366
FT ACT_SITE 322 323
FT DISULFID 312 362
FT DISULFID 321 345
FT DISULFID 337 358
FT DOMAIN 218 229
FT CARBOHYD 628 628
FT VARSPIC 311 365
FT VARSPIC 616 627
FT CONFLICT 575 577
SQ SEQUENCE 765 AA; 86882 MW; CF51FCCCE305A0CF CRC64;

Query Match 48.6%; Score 1971.5; DB 1; Length 765;
Best Local Similarity 49.7%; Pred. No. 1.2e-95;
Matches 401; Conservative 134; Mismatches 173; Indels 99; Gaps 20;

QY 5 LAILLAAATARALEV-----PTDGNAG---LLAEPOIAMFCGRLLNMNMVQNGKWDSP 56
DB 15 LVLLGLGLTAPAAALAGYIEALANAGTGFVAEPQIAMFCGKLNHMHVNIQTGWEPDP 74
QY 57 SGRKCTIDTREGILQCYQYPELOITNVVEANQPVYTIQNWCKRGRKQCKTHPHFVPIYR 116
DB 75 TGTGKSLGTEKEVLQCYQYPELOITNVWEANQPVNIDSWCRDRKQCRS--HIVIPK 132
QY 117 CLVGFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLDHGMLLPCGI 176
DB 133 CLVGFVSDVLLVPENCQFFHQRMEVCCKHQRWHTVVKAECLTEGMLTLYSGMLLPCGV 192
QY 177 DKFRGVEFVCCPLAE--ESDNVDSADAEEDSDVWVGADTDYA-DGSEDKVVEVAEEEE 233
DB 193 DQFHGTGYVCCPTKVVDSDTMSKKEEEEBE---DEEDYALDKSEFPTAELEDT 248
QY 234 VAEEVEEADDDDDGDEVEEEAEEPEE-----ATERITTSIATTTTTTTSVEEVV 287
DB 249 EAAADEDEDEEEEEEVEEDRDYVYDFKGDYNEENPEPSDDGTISDKETAHDV 308
QY 288 REVCSQAEQTPGRAMISRWFDVTGKCAPFFYGGCGGNRNNDTEECMAVCSAMSQ 347
DB 309 KAVCSQAEAMTPGRVAMPWRPFDLSKGVKRVFTYGGCGGNRNNFSEDCYMAVCKTMI-- 366
QY 348 SLKTTQEPPLARDPVKLPTAATSTPDVADVKYLETGCDENEHAFOKAKERLEAKHERMS 407
DB 367 -----PPTPEPT-----NDVDVYFETSADDNEHAFOKAQEIRHRSMD 408
QY 408 QVREWEAEQAQKLPKADKKAVIOHFQEKVESLQEAANEERQQLVETHMARVEAMLND 467
DB 409 RVKKEWEAEQAQKLPKAEQRTQLHFQAMVKALEKAASEKQQLVETHLARVEAMLND 468
QY 468 RRLALLENITAOQVPPRRHVFNMKLYVRAEQKDRQHTLKHFEHVRWMDPKKAAQIR 527

```

```

DB 469 RRIALENYLAALQSDPPRPHRILOALRRYVRAENKORLHTIRHYQHVLAVDPEKAAQMK 528
QY 528 SQVMTHLRVIVERNQSLSLLYNYPVAEETODEVDELLOKEQNYSDVLANWLSEPRIS 587
DB 529 SQVMTHLRVIEERNQSLSLLYKVPYVAQEIQEIDELOQOR-----ADM----- 574
QY 588 YGNDALMPSLTETKTIVELLPVNGEFSUDDLPQWHSFGADSVDPANTENEVEPVDARPAAD 647
DB 575 ---DQFTSSISENPVDVR---VSSEES-EEIPPPFPF--HFPFSLSENE----DTQPELY 621
QY 648 RGLTTRPGSGLTN-----IKTEE---ISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGS-- 697
DB 622 HPM--KKGSGMAEQDGGGLIGAEKVKNSKMDENMVIDELDV--KEMIFNAERVGGL 677
QY 698 -----NKGAIIGLWVGGVVIATVIVITLVMLKKKQYTSIHIGVVEVDAV 742
DB 678 EEPDSVGPLRDFSLSSSALIGLLVIAVIAIVISLVMLRKROYGTISHGIVEVHPML 737
QY 743 TPEERHLSKMQQNGYENPTYKFFEQMQ 769
DB 738 TPEERHLNKMONGYENPTYKYLEQMQ 764

RESULT 7
APP2_MOUSE
ID APP2_MOUSE STANDARD; PRT; 695 AA.
AC O06335;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Amyloid-like protein 2 precursor (CDEI-box binding protein) (CDEBP).
GN APLP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA von der Kammer H.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-246 FROM N.A.
RX MEDLINE=94032480; PubMed=8218408;
RA Hanes J., von der Kammer H., Kristjansson G.I., Scheit K.H.;
RT "The complete cDNA coding sequence for the mouse CDEI binding
protein.";
RL Biochim. Biophys. Acta 1216:154-156(1993).
[3]
RP SEQUENCE OF 185-695 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Heart;
RX MEDLINE=93129193; PubMed=1482349;
RA Vidal F., Blangy A., Rassoulzadegan M., Cuzin F.;
RT "A murine sequence-specific DNA binding protein shows extensive local
similarities to the amyloid precursor protein.";
RL Biochem. Biophys. Res. Commun. 189:1336-1341(1992).
[4]
RP SEQUENCE OF 1-35 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=96029629; PubMed=7592716;
RA von Koch C.S., Lahiri D.K., Mammen A.L., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Sisodia S.S.;
RT "The mouse APLP2 gene. Chromosomal localization and promoter
characterization.";
RL J. Biol. Chem. 270:25475-25480(1995).
CC -!- FUNCTION: BINDS TO THE DNA 5'-GTCAATG-3' (CDEI BOX) WHICH PLAYS
AN IMPORTANT ROLE IN THE EARLY DEVELOPMENT OF EMBRYOS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN AND NUCLEAR
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; 222592; CAA80306.1; -
 CC EMBL; M97216; AAA20039.1; -
 CC EMBL; U34291; AAC52318.1; -
 CC PIR; JC1404; JC1404.
 CC HSSP; P05067; IMWP.
 CC MGD; MGI:88047; APLP2.
 CC InterPro; IPR001868; A4_APP.
 CC Pfam; PF02177; A4_EXTRA; 1.
 CC PRINTS; PR00203; AMYLOIDA4.
 CC SMART; SM00006; A4_EXTRA; 1.
 CC PROSITE; PS00319; A4_EXTRA; 1.
 CC PROSITE; PS00320; A4_INTRA; 1.
 CC Transmembrane; DNA-binding; Signal; Nuclear protein.
 CC SIGNAL 1 29
 CC CHAIN 30 695
 CC DOMAIN 30 624
 CC TRANSMEM 625 648
 CC TRANSMEM 649 695
 CC DOMAIN 218 294
 CC DOMAIN 218 231
 CC DOMAIN 256 266
 CC CARBOHYD 485 485
 CC CONFLICT 185 189
 CC SEQUENCE 695 AA; 78944 MW; BBF4B95AAB2A0311 CRC64;

Query Match 42.0%; Score 1704.5; DB 1; Length 695;
 Best Local Similarity 46.9%; Pred. No. 8.5e-82;
 Matches 371; Conservative 113; Mismatches 170; Indels 137; Gaps 24;

QY 5 LALLLLAAMTARALEV-----PTDGNAG---LLAEPOIAMFCGRGLNMHNVQNGKWDSP 56
 DB 15 LVLLLLGLTAPAAALAGYTEALANAGTGFAVEPOIAMLCGLNMHNVQNGKWDSP 74
 QY 57 SGTKCTIDTREGILOQCEVYVPELQITNVVNEANQPTVQNWCKRGRKQCKTHPHFVYPR 116
 DB 75 TGTKSCIGTKEEVLQCEIYVPELQITNVVNEANQPTVQNWCKRGRKQCKTHPHFVYPR 132
 QY 117 CLNGFEVSDALLVPDKCFPLHQRMDVCEHLLHWHVAVETCSEKSTNLHDYGMLLPCGI 176
 DB 133 CLNGFEVSDVLLVPDNCQFPHQRMVECEKHQRHWTLVKBEACLTGLTLYSGMLLPCGV 192
 QY 177 DKFRGFEVFCPLAEESDNVSDADAEEDSDVVMGGADTDYADGSDKVVVEVAEEVAE 236
 DB 193 DQFHGTGYVCCP---QTKTVDS-----DSTMSK-----EEEE--- 221
 QY 237 VEEDAEADDDDEDEGEVEEAEPEEATERTSTATTTTTESEVEEVVREVCSEAE 296
 DB 222 -EBEDEDEDEYDLQSEPTTEADLEDETE---AAADEEEDEEGEREVED----- 270
 QY 297 TGPCRAMISRWYEDVTEGKCAPFFYCGCGNRNFDTEECYCAVCSAMSQSLLKTTQEP 356
 DB 271 -----RDYYD-----PF-----KGDYNEEPT-----PSEGTI---SDKE 301
 QY 357 LARDPKLPPTAATPDVADKYLETGPDENEHAFKAKERLEAKHRMSQVMREWEA 416
 DB 302 IVHD-VKVPPTPLTND-VDVILETSADNEHAFKAKERLEAKHRMSQVMREWEA 359
 QY 417 EROAKNLPKADKKAVIOHQFEQVYESLEQEAEANERQOLVETHMARVEAMLNRRRLALENY 476
 DB 360 ELQAKNLPKTERQTLQHQFQAVKALEKAASEKQQLVETHLARVEAMLNRRRLALENY 419
 QY 477 ITALQAVPPRPFRVFNKLVYVRAEQDKRQHTLKHFEHVMVDPKKAAQIRSQVMTHLRV 536
 DB 420 LAALQSDPPRPHRILOAKRYVRAENKRLHTIRHQVHLAVDPKAAQKMSQVMTHLRV 479
 QY 537 IYERMNQSLLYNVPAVAEEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPS 596

DB 480 IEERNQSUULLYKYPYVAQEIQEIDELLQEQR-----ADM-----DQFTSS 522
 QY 597 LRETETTVLLPVNGEFSLDDLPWHSFGADSVPAANTENEVEPDARPAADRLTTRPGS 656
 DB 523 ISENPDVVRVSESE-EIPPPHPLHPF-----PSLSENE-----GSGMAEQDG- 565
 QY 657 GLTNKTEBI-SEVKMDAEFRHDSGYEVHVKLVFAEDVGS-----N 698
 DB 566 GLIGAEKVKINSKMDENMVIDETLDV--KEMIFNAERVGLGEPEESVGLREDEFSLS 623
 QY 699 KGAIIGLMVGGVVIATVITLVMLKKQYTSIHGGVVEVDAVTPERHLSKMQQNGYE 758
 DB 624 SNALIGLVIAVIAIVIVISLVMLRKKQYGTISHGIVEVDPMLTPERHLNKMNHGTE 683
 QY 759 NPTYKFEQMQ 769
 DB 684 NPTYKYLEQM 694

RESULT 8

APPL_HUMAN
 ID APPL_HUMAN STANDARD; PRT; 650 AA.
 AC P51693; C00113;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Amyloid-like protein 1 precursor (APLP).
 GN APLP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808960; PubMed=9428684;
 RA Paliga K., Peraus G., Kregar S., Duwrrwang U., Hesse L., Muthaup G.,
 RA Masters C.L., Beyreuther K., Weidemann A.;
 RT "Human amyloid precursor-like protein 1--cDNA cloning, ectopic
 RT expression in COS-7 cells and identification of soluble forms in the
 RT cerebrospinal fluid.";
 RL Eur. J. Biochem. 250:354-363(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9818087; PubMed=9521588;
 RA Lenkkeri U., Kestila M., Lamerdin J., McCready P., Adamson A.,
 RA Olsen A., Tryggvason K.;
 RT "Structure of the human amyloid-precursor-like protein gene APLP1 at
 RT 19q13.1.";
 RL Hum. Genet. 102:192-196(1998).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. C-TERMINALLY
 CC PROCESSED IN THE GOLGI COMPLEX AND IS THEN SECRETED.
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN BRAIN, PARTICULARLY
 CC IN THE CEREBRAL CORTEX POSTSYNAPTIC DENSITY.
 CC -!- PTM: N- AND O-GLYCOSYLATED.
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U48437; AAB96331.1; -
 CC EMBL; AD000864; AAB50173.1; -
 CC HSP; P05067; IMWP.
 CC Genew; HGNC:597; APLP1.
 CC MIM; 104775; -
 CC InterPro; IPR001868; A4_APP.
 CC Pfam; PF02177; A4_EXTRA; 1.
 CC PRINTS; PR00203; AMYLOIDA4.


```
DR SMART; SM00006; A4_EXTRA; 1.  
DR PROSITE; PS00319; A4_EXTRA; 1.  
DR PROSITE; PS00320; A4_INTRA; 1.  
KW Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 650  
FT DOMAIN 22 580  
FT TRANSMEM 581 603  
FT DOMAIN 604 650  
FT DOMAIN 640 643  
FT DOMAIN 241 247  
FT DOMAIN 264 268  
FT CARBOHYD 337 337  
FT CARBOHYD 461 461  
FT CARBOHYD 551 551  
FT CONFLICT 48 48  
FT SEQUENCE 650 AA; 72202 MW; 121A034B708C67CA CRC64;  
  
Query Match 28.5%; Score 1150.5; DB 1; Length 650;  
Best Local Similarity 35.7%; Pred. No. 2.9e-53;  
Matches 279; Conservative 114; Mismatches 218; Indels 171; Gaps 20;  
  
QY 1 MLPGLALLLL-----LAATARALETPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKW 52  
DB 23 LLPULLLLRAQPAIGSLAGSGAPEAP--GSA-----QVAGLCGRLLTLHRLDRTGRW 74  
QY 53 DSDPSGFKTCIDTKEGILQYQCVPELQITNVVEANQPVTIQNWCKRGRKCKTHPHF- 111  
DB 75 EPDPQSRCLRDQPVLEVCROMYELQIARVEQATQAIPIWRCGSGSGSCAHPHQ 134  
QY 112 VIPYRCLVGFVSADLLVPDKKFLQHOERMDVCETHLHWHVTAKECTSEKSTNLHDYGL 171  
DB 135 WPPRCLPGFVSALLVPGRFLQHOERMDQCESSTRRHOEAQACSSQGLLHSGGML 194  
QY 172 LPGAIDFRGVFVCCPLAESDNVDSADAEEDSDVWVGADTDVADGSEKVVVEAE 231  
DB 195 LPGAIDFRGVFVCCPLAESDNVDSADAEEDSDVWVGADTDVADGSEKVVVEAE 231  
QY 232 EEVAEEBEEADDEDEDEDEVEEAEEPEEATERTTSIATTTTTTESVEVVREVC 291  
DB 243 EE-----EEESFPQPVDDYFVEPPQAESEE----- 267  
QY 292 SEQAETGCRAMISRYFVTEGKCAPFFYGGCGGNRNFDETEYCMVCGSAMSLSLK 351  
DB 268 -----ETVP-----PPSSHTLAVVG-----K 283  
QY 352 TTQEPLEARDPVKLTAASTPDADVLYLETPGDENEHAHQAKERLEAKHRERMSQVR 411  
DB 284 VT-----PTPRT-----DGVDFYFGMPGEISSEHGFELRAKMDLEERRMQINEVMR 330  
QY 412 EWEEAERQAKNLKADKAKVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRL 471  
DB 331 EWAMADNQSNLKPADQALNEHFQSTLQLEQVSGERQLVETHATRVIALINDQRR 390  
QY 472 ALENYITAAQVPPRHVFNMLKVVYRAEQKDRQHTLKHFEHVRVMDPKAAQIRSQVM 531  
DB 391 ALGEFLAALQADPQABRVLLALRYLRABEQEQRUTLRHYVAAVADPEKAQOMRFQVH 450  
QY 532 THLRVIVERNQSLSLLYNPVAAEETODEVDELLOKEQNSDDVLNMISEPRISYEND 591  
DB 451 THLQVIEERNQSLGLLDQNPFLAQLERLQIOELLSEH----- 489  
QY 592 ALMPSLTETKTVLLPVNGEFLSD--DLQPHWSFGADSVPAANTEVEPVDARPAADRG 649  
DB 490 -LGPSELEA-----PAPGSSDEKGLQPPDS--KDOTPM-----TLPKG 526  
QY 650 LTRPGSGLNITKEETISEYKMDAEFRHDSGYEVHH---QKLVFFAEDVGSNKGAIIGLM 706  
DB 527 STEQDAASPEKEMNPLEQYERKVNASVPRGPFPHSSEIQRDLEAPAGTGVSRVAVSGLL 586  
QY 707 VGGVVIATVITLML-KKKQVTSHHGVVEVDAVTPPEERLSKMQONGYENPTYKFF 765  
DB 587 IMGAGGSLIVLSMLLRKKPKYGAISHGVVEVDPMLTEEQQLRELQHRGYNPTYREL 646
```

QY 766 EQ 767
DB 647 EE 648

RESULT 9
APPL_MOUSE
ID APPL_MOUSE STANDARD; PRT; 653 AA.
AC Q03157;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Amyloid-like protein 1 precursor (APLP).
GN APLP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93066322; PubMed=1279693;
RA Wasco W., Bupp K., Magendantz M., Gusella J.F., Tanzi R.E.,
RA Solomon F.;
RT "Identification of a mouse brain cDNA that encodes a protein related
to the Alzheimer disease-associated amyloid beta protein precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10758-10762(1992).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MAY BE PROCESSED
IN THE GOLGI COMPLEX; HOWEVER, IT IS NOT YET CLEAR WHETHER APLP
IS SECRETED.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L04538; AAA37247.1; -;
CC PIR; A46362; A46362.
CC HSP; P05067; IMWP.
CC MGD; MGI:88046; APLP1.
CC InterPro; IPR001868; A4 App.
CC Pfam; PF02177; A4_EXTRA; 1.
CC PRINTS; PR00203; AMYLOIDA4.
CC SMART; SM00006; A4_EXTRA; 1.
CC PROSITE; PS00319; A4_EXTRA; 1.
CC PROSITE; PS00320; A4_INTRA; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 20
FT CHAIN 21 653
FT DOMAIN 21 583
FT DOMAIN 584 606
FT TRANSMEM 607 653
FT DOMAIN 607 653
FT DOMAIN 643 646
FT DOMAIN 263 271
FT DOMAIN 464 464
FT CARBOHYD 554 554
FT CARBOHYD 554 554
FT SEQUENCE 653 AA; 72751 MW; 56516DC3EA40E4B0 CRC64;

Query Match 28.4%; Score 1150.5; DB 1; Length 653;
Best Local Similarity 35.4%; Pred. No. 5.3e-53;
Matches 274; Conservative 120; Mismatches 228; Indels 153; Gaps 19;

QY 1 MLPGLALLLLAAWTARA-LEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSPSGT 59
DB 22 LLP-LSLLLLRAQLAVGNLVGSPSAEAPGSAQVAGLCGRLLTLHRLDRTGRWEPDPQS 80
QY 60 KTCIDTKEGILQYQCVPELQITNVVEANQPVTIQNWCKRGRKCKTHPHF-VIPYRCL 118
DB 60 KTCIDTKEGILQYQCVPELQITNVVEANQPVTIQNWCKRGRKCKTHPHF-VIPYRCL 118

Db 496 KREAMTCVTOALTEQPPNAHVCKLQKLLRALHKLORAHALAHYRHLNSGGPGGLEAAA 555
 QY 525 QIRSOVTHLRVIERWNOSLSLLYNNPVAEEI-----QDEV----- 562
 Db 556 SERPTLERLIDIDRAVNSQMTMLKRYPELSAKIAQLMNDYIILALRSKDDIPGSSGLMSE 615
 QY 563 -----DELLOKQBNYSDDVLAN 579
 Db 616 EAEAGILDKYRVEIERKVAERLRLAEKQKQRAERKLEKRLAEAKKVDMDLKS 675
 QY 580 MISE-----PRISYGNDAIM-----PSLTETKTTVELLPVNG 611
 Db 676 QVABEQSQPTQSSQSOAQOQOQEKSLPGKELGPDAAALVTAANPNLETKS----- 726
 QY 612 EFSLDDLPQWHSFGADSVPANTENEVEVDPAADARGLTRPGSGLTNKTKEISEVKM 671
 Db 727 EKLSLDE-----YGEATVSTTKVQTVLPTVDDDAVQRAVEDVAAA-----VAHQEA 773
 QY 672 DAEPHDGSGYEVHKKLVF-----FAEDVGSNK-----GAIIGLMVGGVVIATVIVITLVL 723
 Db 774 EPQVQHFMTDHLGHRSSFSLSRLREFAQHAHAKEGRNVYFTLSFAGIALMAAEVFGVAVA 833
 QY 724 KKKQYTSIH-HGVVEVDAAVTP-----EERHLSKMOQNGYENPTYKFFE 766
 Db 834 KWRTSRSPHAQGFTEVDQNVTHHPVIREKIVPNMQINGENPTYKYFE 883

RESULT 12

A4_MACFA STANDARD; PRT; 87 AA.
 ID A4_MACFA
 AC P53601;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein (Fragment).
 GN APP.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL MEDLINE=91273117; PubMed=1905108;
 RA Podlinsky M.B., Tolan D.R., Selkoe D.J.;
 RT "Homology of the amyloid beta protein precursor in monkey and human
 RT supports a primate model for beta amyloidosis in Alzheimer's
 RT disease";
 RT Am. J. Pathol. 138:1423-1435(1991).
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M58726; AAA36828.1; -;
 CC HSP; P05067; IAAAP
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR002223; Kunitz_BPTI.
 CC Pfam; PF00014; Kunitz_BPTI; 1.

DR PRINTS: PR00759; BASICPTASE.
 DR ProDom: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00131; KU; 1.
 DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
 DR PROSITE: PS00320; A4_INTRA; PARTIAL.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS00379; BPTI_KUNITZ_2; 1.
 KW Glycoprotein; Amyloid; Neurone; Alternative splicing;
 KW Serine protease inhibitor.
 FT NON_TER 1
 FT DOMAIN 4 79 BPTI/KUNITZ INHIBITOR.
 FT ACT_SITE 16 17 REACTIVE BOND.
 FT DISULFID 6 56 BY SIMILARITY.
 FT DISULFID 15 39 BY SIMILARITY.
 FT DISULFID 21 52 BY SIMILARITY.
 FT NON_TER 87 87
 SQ SEQUENCE 87 AA; 9608 MW; B67CG90DEE0EE7FF CRC64;
 Query Match 11.3%; Score 457; DB 1; Length 87;
 Best Local Similarity 95.4%; Pred No. 5, 6e-18;
 Matches 83; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 286 VREVCEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEYCMAYCGSAM 345
 Db 1 VREVCEQAETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNFDTEYCMAYCGSVN 60
 QY 346 SQSLKTTQEPPLARDPVKLPPTAASTP 372
 Db 61 QSLSRKTREPLTRDPVKLPPTAASTP 87
 RESULT 13
 A4_MACMU
 ID A4_MACMU
 AC P29216;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alzheimer's disease amyloid A4 protein (fragment).
 GN APP.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Koo E.H., Sisodia S.S., Price D.L.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
 CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
 CC G(O) (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; APP(395), APP(563), APP(695),
 CC APP(751) AND APP(770) (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15985; CAA34116.1; -;
 CC FIR; S06678; S06678.
 CC HSP; P05067; IAAAP.
 CC InterPro; IPR001868; A4_APP.
 CC InterPro; IPR002223; Kunitz_BPTI.


```

DR Pfam: PF00014; Kunitz_BPTI; 1.
DR ProDom: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
DR Glycoprotein; Amyloid; Neurone; Alternative splicing;
KW Serine protease inhibitor.
FT NON_TER 1 1
FT DOMAIN 1 76 BPTI/KUNITZ INHIBITOR.
FT ACT_SITE 13 14 REACTIVE BOND.
FT DISULFID 3 53 BY SIMILARITY.
FT DISULFID 12 36 BY SIMILARITY.
FT DISULFID 28 49 BY SIMILARITY.
FT NON_TER 76 76
FT SEQUENCE 76 AA; 8527 MW; 492BF3069AB082A1 CRC64;

Query Match 9.9%; Score 403; DB 1; Length 76;
Best Local Similarity 94.7%; Pred. No. 3.1e-15;
Matches 72; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 289 EVCSQETGCRAMISRWFVTEGKCAPFFYGGGNGNNFDTBEYCMVCGSAMQS 348
DB 1 EVCSQETGCRAMISRWFVTEGKCAPFFYGGGNGNNFDTBEYCMVCGSAMQS 60
QY 349 LLKTTQEPRLARDPKVL 364
DB 61 LRKTTREPLTRDPVKL 76

RESULT 14
A4_BOVIN
ID A4_BOVIN STANDARD; PRT; 59 AA.
AC Q28033;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
DE APP.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OX Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56124; CAA39589.1; -
CC DR EMBL; X56126; CAA39591.1; -
CC DR HSSP; P05067; IBA4.
CC InterPro: IPR001868; A4_APP.

```

```

DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF03494; Beta-APP; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein; Amyloid; Neurone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 7 49 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 34 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 35 58 POTENTIAL.
FT DOMAIN 59 >59 CYTOPLASMIC (POTENTIAL).
FT NON_TER 59 59
FT SEQUENCE 59 AA; 6414 MW; F43469D488A2E12D CRC64;

Query Match 7.2%; Score 292; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

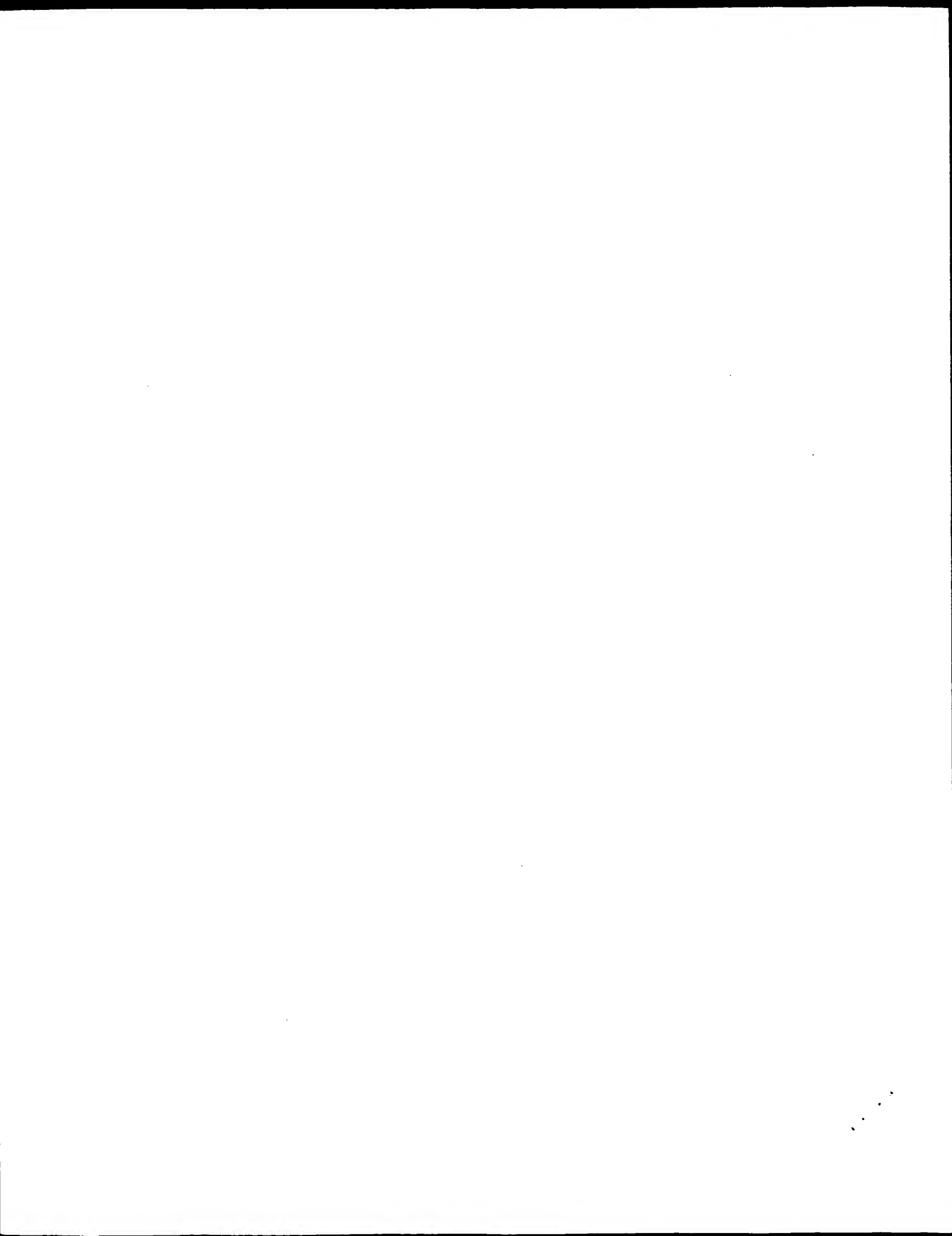
QY 666 ISEVKMDAEFRHDSGYEVHQLVFFAEVDGSGNKGAIIGLMVGGVVIVITLVMLK 724
DB 1 ISEVKMDAEFRHDSGYEVHQLVFFAEVDGSGNKGAIIGLMVGGVVIVITLVMLK 59

RESULT 15
A4_RABIT
ID A4_RABIT STANDARD; PRT; 58 AA.
AC Q28748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog [Contains: Beta-amyloid
DE protein (Beta-APP) (A-beta)] (Fragment).
DE APP.
GN Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92017079; PubMed=1656157;
RA Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT "Conservation of the sequence of the Alzheimer's disease amyloid
RT peptide in dog, polar bear and five other mammals by cross-species
RT polymerase chain reaction analysis.";
RL Brain Res. Mol. Brain Res. 10:299-305(1991).
CC -!- FUNCTION: FUNCTIONAL NEURONAL RECEPTOR WHICH COUPLES TO
CC INTRACELLULAR SIGNALING PATHWAY THROUGH THE GTP-BINDING PROTEIN
CC G(O) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE APP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56129; CAA39594.1; -
CC DR HSSP; P05067; IBA4.
CC InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF03494; Beta-APP; 1.
DR PROSITE: PS00319; A4_EXTRA; PARTIAL.
DR PROSITE: PS00320; A4_INTRA; PARTIAL.
KW Glycoprotein; Amyloid; Neurone; Transmembrane.
FT NON_TER 1 1
FT CHAIN 6 48 BETA-AMYLOID PROTEIN (POTENTIAL).
FT DOMAIN <1 33 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 34 57 POTENTIAL.
FT DOMAIN 58 >58 CYTOPLASMIC (POTENTIAL).
FT NON_TER 58 58

```


SQ SEQUENCE 58 AA; 6300 MW; F434209D88EBA82D CRC64;
 Query Match 7.1%; Score 288; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 2.1e-09;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 667 SEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLWVGWVIATVIVITLVMLK 724
 Db 1 SEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLWVGWVIATVIVITLVMLK 58

Search completed: January 28, 2003, 14:14:59
 Job time : 15.7782 secs



GenCore version 5.1.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 38.9101 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLAATAWTALEV.....KMQQNGYENPTYKFFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-rodent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*
15: sp-rvivirus.*
16: sp-bacteriap.*
17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3981	98.1	770	6 Q9TUI0	Q9tui0 sus scrofa
2	3719.5	91.7	751	13 Q9DGJ7	Q9dgj7 gallus gall
3	3590.5	88.5	695	6 Q95KN7	Q95kn7 macaca fasc
4	3522.5	86.8	695	11 Q60496	Q60496 cavia sp. p
5	3489.5	86.0	695	11 P97487	P97487 mus musculus
6	3403.5	83.9	747	13 Q91963	Q91963 xenopus ap
7	3377.5	83.2	695	13 Q9DGJ8	Q9dgj8 gallus gall
8	3163.5	78.0	693	13 Q98SG0	Q98sg0 xenopus lae
9	3152.5	77.7	695	13 Q98SF9	Q98sf9 xenopus lae
10	3065	75.5	607	11 Q99K32	Q99k32 mus musculus
11	2914	71.8	699	13 Q57394	Q57394 narke japon
12	2831	69.8	547	4 Q13764	Q13764 homo sapien
13	2785.5	68.6	738	13 Q90W28	Q90w28 brachydanio
14	2763.5	68.1	780	13 Q73683	Q73683 tetraodon f
15	2717	67.0	569	13 Q9PVL1	Q9pvl1 gallus gall
16	2699	66.5	737	13 Q93279	Q93279 fugu rubrip

17	2562.5	63.1	534	13	Q93296	Q93296 gallus gall
18	2446.5	60.3	484	4	Q13793	Q13793 homo sapien
19	2437	60.1	694	13	Q8UUR9	Q8uur9 brachydanio
20	2288.5	56.4	612	13	Q919E7	Q919e7 brachydanio
21	1960.5	48.3	763	11	Q61482	Q61482 mus musculus
22	1956.5	48.2	751	11	Q60709	Q60709 mus musculus
23	1733.5	42.7	695	4	Q13861	Q13861 homo sapien
24	1721	42.4	669	4	Q14662	Q14662 homo sapien
25	1709.5	42.1	695	11	Q64348	Q64348 mus musculus
26	1604.5	39.5	472	13	Q8UUS0	Q8uus0 brachydanio
27	1569	38.7	305	4	Q9BT38	Q9bt38 homo sapien
28	1350.5	33.3	357	13	Q8UUI8	Q8uii8 brachydanio
29	1302	32.1	523	4	Q14594	Q14594 homo sapien
30	1267	31.2	522	4	Q9BT36	Q9bt36 homo sapien
31	1159.5	28.6	650	4	Q96A32	Q96a32 homo sapien
32	1150.5	28.4	654	11	Q8VC38	Q8vc38 mus musculus
33	788	19.4	160	11	Q9Q278	Q9qz78 cavia sp. p
34	771	19.0	239	13	Q8UUI7	Q8uii7 brachydanio
35	739	18.2	887	5	Q9U4H3	Q9u4h3 drosophila
36	738	18.2	887	5	Q9TVU0	Q9tvu0 drosophila
37	678	16.7	136	6	P79307	P79307 sus scrofa
38	661	16.3	816	5	Q9W5F1	Q9w5f1 drosophila
39	561	13.8	182	11	Q9CYS4	Q9cys4 mus musculus
40	493	12.1	97	4	Q13778	Q13778 homo sapien
41	478	11.8	97	6	Q28673	Q28673 oryctolagus
42	411	10.1	82	4	P78438	P78438 homo sapien
43	402	9.9	75	4	Q9UQ58	Q9uq58 homo sapien
44	393.5	9.7	82	4	Q16019	Q16019 homo sapien
45	389.5	9.6	82	4	Q16014	Q16014 homo sapien

ALIGNMENTS

RESULT 1

Q9TUI0 PRELIMINARY; PRT; 770 AA.
ID Q9TUI0
AC Q9TUI0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Anyloid precursor protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura A., Takahashi T.;
RT "Anyloid Precursor Protein 770."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB032550; BAA84580.1; -
DR HSSP; P05067; IAAP.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR InterPro; IPR002223; Kunitz_BPTI.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR Pfam; PF00014; Kunitz_BPTI; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR PRINTS; PR00759; BASICPTASE.
DR ProDom; PD000222; Kunitz_BPTI; 1.
DR SMART; SM00006; A4_EXTRA; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
SQ SEQUENCE 770 AA; 86961 MW; 5F7A1DCB2BCC583E CRC64;
Query Match 98.1%; Score 3981; DB 6; Length 770;
Best Local Similarity 97.8%; Pred. No. 6.1e-232;

Matches 753; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSDPSGTK 60
 Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSDPSGTK 60
 Qy 61 TCIDTKGILQYCOEYVPELQITNVVEANQPVTIQNCKRGKCKOCTHPIHPIYRCVLG 120
 Db 61 TCIDTKGILQYCOEYVPELQITNVVEANQPVTIQNCKRGKCKOCTHPIHPIYRCVLG 120
 Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Qy 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEAEVEAEVEE 240
 Db 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEAEVEAEVEE 240
 Qy 241 EADDEDEDEDEVEEAEEPEEATERTTSIATTTTTTTESEVEEVREVCSQAETGPC 300
 Db 241 EADDEDEDEDEVEEAEEPEEATERTTSIATTTTTTTESEVEEVREVCSQAETGPC 300
 Qy 301 RAMISRWYFDVTGKCAPFYGGCGGNRNDFTEECMAVCGSAMQSLLKTTQEPFLARD 360
 Db 301 RAMISRWYFDVTGKCAPFYGGCGGNRNDFTEECMAVCGSAMQSLLKTTQEPFLARD 360
 Qy 361 PVKLPKADKAVIQHFOEKVESLEQEAEANERQQLVETHMARVEAMLNDRRLALENIT 480
 Db 361 PVKLPKADKAVIQHFOEKVESLEQEAEANERQQLVETHMARVEAMLNDRRLALENIT 480
 Qy 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 Db 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
 Qy 541 MNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
 Db 541 MNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
 Qy 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGTLTPRPGSLTN 660
 Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGTLTPRPGSLTN 660
 Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPFAEDVGSNKGAIIGLMVGGVVIATVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFPFAEDVGSNKGAIIGLMVGGVVIATVITL 720
 Qy 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770
 Db 721 VMLKKKQYTSIHGVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 2
 Q9DGJ7 PRELIMINARY; PRT; 751 AA.
 AC Q9DGJ7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE Beta-amyloid precursor protein 751 isoform.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sarasa M., Rodolose A., Sorribas V.;
 RT "Cloning of full-length chicken beta-amyloid precursor protein
 isoforms.";

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF289219; AAG00594.1; -.
 DR HSSP: P05067; 1BA4.
 DR InterPro: IPR001868; A4_APP.
 DR InterPro: IPR001255; Beta-APP.
 DR InterPro: IPR002223; Kunitz_BPTI.
 DR Pfam: PF02177; A4_EXTRA; 1.
 DR Pfam: PF03494; Beta-APP; 1.
 DR Pfam: PF00014; Kunitz_BPTI; 1.
 DR PRINTS: PR00203; ANYLOIDA4.
 DR PRINTS: PR00759; BASICPTASE.
 DR PRODom: PD000222; Kunitz_BPTI; 1.
 DR SMART: SM00006; A4_EXTRA; 1.
 DR SMART: SM00131; KU; 1.
 DR PROSITE: PS00319; A4_EXTRA; 1.
 DR PROSITE: PS00320; A4_INTRA; 1.
 DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
 DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
 KW Serine protease inhibitor.
 SQ SEQUENCE 751 AA; 84705 MW; E78E9413A8033D84 CRC64;

Query Match 91.7%; Score 3719.5; DB 13; Length 751;
 Best Local Similarity 91.7%; Pred. No. 3.5e-216;
 Matches 708; Conservative 18; Mismatches 23; Indels 23; Gaps 4;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSDPSGTK 60
 Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSDPSGTK 60
 Qy 61 TCIDTKGILQYCOEYVPELQITNVVEANQPVTIQNCKRGKCKOCTHPIHPIYRCVLG 120
 Db 61 TCIDTKGILQYCOEYVPELQITNVVEANQPVTIQNCKRGKCKOCTHPIHPIYRCVLG 120
 Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Qy 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEAEVEAEVEE 238
 Db 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEAEVEAEVEE 240
 Qy 239 EEEADDEDEDEDEVEEAEEPEEATERTTSIATTTTTTTESEVEEVREVCSQAETG 298
 Db 241 DEEADDD-DEDDGDI-BETEEYEAEATERTTSIATTTTTTTESEVEEVREVCSQAETG 298
 Qy 299 PCRAMISRWYFDVTGKCAPFYGGCGGNRNDFTEECMAVCGSAMQSLLKTTQEPPLA 358
 Db 299 PCRAMISRWYFDVTGKCAPFYGGCGGNRNDFTEECMAVCGSV----- 344
 Qy 359 RDPVKLPTTAASTPDVADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMEWEAE 418
 Db 345 -----LPTTAASTPDVADVKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMEWEAE 399
 Qy 419 QAKNLPKADKAVIQHFOEKVESLEQEAEANERQQLVETHMARVEAMLNDRRLALENIT 478
 Db 400 QAKNLPKADKAVIQHFOEKVESLEQEAEANERQQLVETHMARVEAMLNDRRLALENIT 459
 Qy 479 ALQAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIY 538
 Db 460 ALQAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIY 519
 Qy 539 ERMNQSLSLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPST 598
 Db 520 ERMNQSLSFLYNNVPAVEEIQDEVDLQKEQNYSDVLANNISPRISYGNDAIMPST 579
 Qy 599 ETKTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGTLTPRPGSL 658
 Db 580 ETKTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGTLTPRPGSL 639
 Qy 659 TNIKTEEISEVKMDAEFRHDSGYEVHHQKLVFPFAEDVGSNKGAIIGLMVGGVVIATVIT 718
 Db 640 TNVKTTEESEVKMDAEFRHDSGYEVHHQKLVFPFAEDVGSNKGAIIGLMVGGVVIATVIT 699

QY 719 TLVLMKKKQYTSIHGVEVDAVTPERHLSKMQONGYENPTYKFEQMON 770
 Db 700 TLVLMKKKQYTSIHGVEVDAVTPERHLSKMQONGYENPTYKFEQMON 751

RESULT 3

Q95KN7 ID Q95KN7 PRELIMINARY; PRT; 695 AA.
 AC Q95KN7:
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Amyloid b-protein precursor.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CEREBELLUM;
 RX MEDLINE=91273117; PubMed=1905108;
 RA Podlinsky M.B., Tolan D.R., Seikoe D.J.;
 RT "Homology of the amyloid beta protein precursor in monkey and human
 RT supports a primate model for beta amyloidosis in Alzheimer's
 RT disease.";
 RL Am. J. Pathol. 138:1423-1435(1991).
 DR EMBL; M58727; AAA36829.1; -;
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PROSITE; PS00319; A4_EXTRA; UNKNOWN_1.
 DR PROSITE; PS00320; A4_INTRA; UNKNOWN_1.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 597 636 POTENTIAL.
 SQ SEQUENCE 695 AA; 78663 MW; 4F6EA0139F969D56 CRC64;

Query Match 88.5%; Score 3590.5; DB 6; Length 695;
 Best Local Similarity 90.1%; Pred. No. 1.9e-208;
 Matches 694; Conservative 1; Mismatches 0; Indels 75; Gaps 1;

QY 1 MLPGLALLLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
 Db 1 MLPGLALLLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
 Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240
 QY 241 EADDEDEDEGDEVEEAEPEYEATERTSTATTTTTTSTESVEEVREVCSQAETGPC 300
 Db 241 EADDEDEDEGDEVEEAEPEYEATERTSTATTTTTTSTESVEEVREVCSQAETGPC 300
 QY 301 RAMISRWYFDVTEGKCAPFFYCGCGGNRNFTDEYCHAVCGSAMSQSLLKTTQEPPLARD 360
 Db 289 ----- 288
 QY 361 PVKLPPTAASDPADVCKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
 Db 289 ---VPTAASDPADVCKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 345
 QY 421 KNLPRADKAVTQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
 Db 346 KNLPRADKAVTQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 405

QY 481 QAVPPRPHVENMLKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVWTHLRVIER 540
 Db 406 QAVPPRPHVENMLKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVWTHLRVIER 465
 QY 541 MNQSLSLLYNVPAAVEETQDEVDDELLOKEQNSDDVLNMISEPRISYGNDAIMPSTLET 600
 Db 466 MNQSLSLLYNVPAAVEETQDEVDDELLOKEQNSDDVLNMISEPRISYGNDAIMPSTLET 525
 QY 601 KTTVELLPVNGEFSLLDQLQPMHSEFADSPANTENEPEVDARPAADRGTLTTRFSGSLTN 660
 Db 526 KTTVELLPVNGEFSLLDQLQPMHSEFADSPANTENEPEVDARPAADRGTLTTRFSGSLTN 585
 QY 661 IKTEISEVKMDAEFRHDSGYEVHVKLVFFAEDVGSNGKGAIIIGLMVGGVVIVITL 720
 Db 586 IKTEISEVKMDAEFRHDSGYEVHVKLVFFAEDVGSNGKGAIIIGLMVGGVVIVITL 645
 QY 721 VMLKKKQYTSIHGVEVDAVTPERHLSKMQONGYENPTYKFEQMON 770
 Db 646 VMLKKKQYTSIHGVEVDAVTPERHLSKMQONGYENPTYKFEQMON 695

RESULT 4

Q60496 ID Q60496 PRELIMINARY; PRT; 695 AA.
 AC Q60496:
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative amyloid precursor protein.
 OS Cavia sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10143;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=97236426; PubMed=9116031;
 RA Beck M., Mueller D., Bigl V.;
 RT "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
 RT alternative splicing.";
 RL Biochim. Biophys. Acta 1351:17-21(1997).
 DR EMBL; X97631; CAA66230.1; -;
 DR HSP; P05067; IBA4
 DR InterPro; IPR001868; A4_APP.
 DR InterPro; IPR001255; Beta-APP.
 DR Pfam; PF02177; A4_EXTRA; 1.
 DR Pfam; PF03494; Beta-APP; 1.
 DR PRINTS; PR00203; AMYLOID4.
 DR SMART; SM00006; A4_EXTRA; 1.
 DR PROSITE; PS00319; A4_EXTRA; 1.
 DR PROSITE; PS00320; A4_INTRA; 1.
 SQ SEQUENCE 695 AA; 78701 MW; 5196A0C4017F16AB CRC64;

Query Match 86.88; Score 3522.5; DB 11; Length 695;
 Best Local Similarity 88.28; Pred. No. 2.4e-204;
 Matches 679; Conservative 7; Mismatches 9; Indels 75; Gaps 1;

QY 1 MLPGLALLLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
 Db 1 MLPGLALLLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNHNHNVQNGKWDSPSGTK 60
 QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
 Db 61 TCIGSGEILQYCEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEDSDVWVGADTDYADGSDKVVEVAEEVAEEVEE 240


```

QY 241 EADDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSESVEEVVVCSEQAETGPC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 EADDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSESVEEVV- 288
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEP LARD 360
Db 289 ----- 288
QY 361 PVKLPPTAASTPDADVCKYLETGPDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 ---VPTTAASTPDADVCKYLETGPDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 345
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 346 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVTHLRVIYER 540
Db 406 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVTHLRVIYER 465
QY 541 MQQSLLLYNVPAAVEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
Db 466 MQQSLLLYNVPAAVEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 525
QY 601 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVPDARPAADRGLTTRPGSGLTN 660
Db 526 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVPDARPAADRGLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 720
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 645
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFEQMQN 770
Db 646 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFEQMQN 695

RESULT 5
P97487
ID P97487 PRELIMINARY; PRT; 695 AA.
AC P97487; P97942;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hippocampal ankyloid protein.
GN APP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAMP8; TISSUE=HIPPOCAMPUS;
RA Flood J.F., Kumar V.B., Sasser T., Word I., Morley J.E.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 581-662 FROM N.A.
RC STRAIN=1295V;
RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecechi M.,
RA Loring J.F., Goate A.M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84012; AAB41502.1; -.
DR EMBL; U82624; AAB40919.1; -.
DR HSSP; P05067; 1MRP.
DR MGD; MGI:88059; App.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00005; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.

```

```

DR PROSITE; PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78414 MW; 9A5FBE2ED261236E CRC64;

Query Match      86.0%; Score 3489.5; DB 11; Length 695;
Best Local Similarity 87.7%; Pred. No. 2.3e-202;
Matches 675; Conservative 6; Mismatches 14; Indels 75; Gaps 1;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAFICGRLLNMHNMVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAFICGRLLNMHNMVQNGKWDSPSGTK 60
QY 61 TCIDTKEGIIQYCOEYVPELQITNVVEANQPVTIQWCKRGRCKTHPHFVTPYRCLVG 120
Db 61 TCIGTKEGIIQYCOEYVPELQITNVVEANQPVTIQWCKRGRCKTHPHFVTPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
QY 181 GFVEVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEAEVEE 240
Db 181 GFVEVCCPLAEESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEAEVEE 240
QY 241 EADDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSESVEEVVVCSEQAETGPC 300
Db 241 EADDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSESVEEVV- 288
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEYCMVCGSAMQSLLKTTQEP LARD 360
Db 289 ----- 288
QY 361 PVKLPPTAASTPDADVCKYLETGPDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 ---VPTTAASTPDADVCKYLETGPDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 345
QY 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 346 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 405
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVTHLRVIYER 540
Db 406 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVTHLRVIYER 465
QY 541 MQQSLLLYNVPAAVEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 600
Db 466 MQQSLLLYNVPAAVEEIQDEVELLQKEQYSDVLANMISEPRISYGNDAIMPSTLET 525
QY 601 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVPDARPAADRGLTTRPGSGLTN 660
Db 526 KTTVELLPVNGEFLDLOPWHSGADSVPAANTENEVEPVPDARPAADRGLTTRPGSGLTN 585
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 720
Db 586 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVITL 645
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFEQMQN 770
Db 646 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFEQMQN 695

RESULT 6
Q91963
ID Q91963 PRELIMINARY; PRT; 747 AA.
AC Q91963;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE APP747.
GN APP747.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae.

```



```

OX NCBI_TaxID=8353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93129227; PubMed=1282805;
RA Okado H., Okamoto H.;
RT "A Xenopus homolog of the human beta-amyloid precursor protein:
RL development regulation of its gene expression.";
RL Biochem. Biophys. Res. Commun. 189:1561-1568(1992).
DR EMBL: S52417; AAB24853.1; -.
DR HSSP: P05067; 1H23.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR InterPro: IPR002223; Kunitz_BPTI.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR Pfam: PF00014; Kunitz_BPTI; 1.
DR PRINTS: PR00203; AMYLOID4.
DR PRINTS: PR00759; BASICPTASE.
DR PRODOM: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00006; A4_EXTRA; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS00279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
SQ SEQUENCE 747 AA; 84893 MW; A75E81885681D948 CRC64;

Query Match      83.9%; Score 3403.5; DB 13; Length 747;
Best Local Similarity 85.2%; Pred. No. 3.9e-197;
Matches 645; Conservative 40; Mismatches 45; Indels 27; Gaps 5;

QY 17 ALEVPDGNAGLAEPOIAMF-CGRLNHMNVQNGKWDSPSGTKTCIDTKEGILQYCOE 75
DB 15 ALEVLVDNGGLAEPOIAMFVARLNHMNVQNGKWDVDSG---CIGTKEGILQYCOE 71

QY 76 VYPELQITNVVEANQPVTVIQNWKCKGRKCKOCTHPHFVPIYRCVLGFEVSDALLVPDKCF 135
DB 72 VYPELQITNVVEANQPVTVIQNWKCKGRKCKOCTHIVVPIYRCVLGFEVSDALLVPDKCF 131

QY 136 LHQERMDVCTHLHWHVAVKETSCKSTNLHDYGMLLPGCIDKFRGVEFVCCPLAESDN 195
DB 132 LHQERMDICETHLHWHVAVKESCKSTNLHDYGMLLPGCIDKFRGVEFVCCPSAESES 191

QY 196 VDSADAEEDSDYVWGADTDYADGSEDKVVEVA---EEVEAVVEEEDDEDDEDGE 253
DB 192 FDSADAEEDCDYVWGADADYVDRSDDKRAVEAQPDVEEVEVEEEDDEDDEDGE 249

QY 254 VESEAEPEYBEATERTTSIATTTTTTSTESVEEYVREVCSEQAETGPCRAMISRWDVTE 313
DB 250 AEEPEPEYBEATERTTSIATTTTTTSTESVEEYVREVCSEQAETGPCRAMISRWDVTE 309

QY 314 GKCAPEFYGGCGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLARDPVKLTPTTAASPD 373
DB 310 SKCAQFYGGCGNRNFDSDCYMAVCGSV-----IPATAASPD 350

QY 374 AVDKYLETGPDENEHAHFQKAKERLEAKHREMSQVWREWEAEERAKNLPKADKAVIQ 433
DB 351 AVDKYLENPNDENHDFLAKERLSEKREKSEVWKEWEAEERAKNLPKADKAVIQ 410

QY 434 HFQEKVESLEQEAERQQLVETHMARVEAMLNDRRLALENYITALQAVPPRPHVFNK 493
DB 411 HFQEKVESLEQEAERQQLVETHMARVEAMLNDRRLALENYITALQADPPRPHVFNK 470

QY 494 LKKYVRAEQKDRQHTLKHFEHVRMVPDKAAQIRSQVWTHLRVIYERMNQSLLYNVA 553
DB 471 LKKYVRAEQKDRQHTLKHFEHVRMVPDKAAQIRSQVWTHLRVINERMNQSFLLYKVA 530

QY 554 VAEIQDEVDLQKQKONTSDVLANNISEPRISYNDALMPSLTETKTVELLPVNGGF 613
DB 531 VAEIQDEVDLQKQKONTSDVLANNISEPRISYNDALMPSLTETKTVELLPVNGGF 590

QY 614 SLDDLQPHWISFGADSPANTENEVEPVDARPAADRLGTLTPGSGGLTNIKTEISEYKMDA 673

```

```

Db 591 NIEDLQPHWISFGADSPANTENEVEPVDARPAADRLGTLTPGSGGLTNIKTEISEYKMD 650
QY 674 EFRHDSGVEVHHQKLVFFAEADVGSNGKAIIGLVGGVVIATVITVLMKKKQYTSIHH 733
Db 651 EYRHTATVEVHHQKLVFFAEADVGSNGKAIIGLVGGVVIATVITVLMKKKQYTIHH 710
QY 734 GYVEVDAAVTPPEERHLSKMQQNGYENPYKPFQEQMN 770
Db 711 GYVEVDAAVTPPEERHLSKMQQNGYENPYKPFQEQMN 747

RESULT 7
QSDGJ8
AC Q9DGJ8 PRELIMINARY; PRT; 695 AA.
ID Q9DGJ8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Beta-amyloid precursor protein 695 isoform.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sarasa M., Rodolosse A., Sorribas V.;
RT "Cloning of full-length chicken beta-amyloid precursor protein
RT isoforms.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AF289218; AAG00593.1; -.
DR HSSP: P05067; 1BA4.
DR InterPro: IPR001868; A4_APP.
DR InterPro: IPR001255; Beta-APP.
DR Pfam: PF02177; A4_EXTRA; 1.
DR Pfam: PF03494; Beta-APP; 1.
DR PRINTS: PR00203; AMYLOID4.
DR SMART: SM00006; A4_EXTRA; 1.
DR PROSITE: PS00319; A4_EXTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
DR PROSITE: PS00320; A4_INTRA; 1.
SQ SEQUENCE 695 AA; 78565 MW; F201ED02AEC86D95 CRC64;

Query Match      83.2%; Score 3377.5; DB 13; Length 695;
Best Local Similarity 84.7%; Pred. No. 1.3e-195;
Matches 654; Conservative 18; Mismatches 21; Indels 79; Gaps 4;

QY 1 MLPLGLALLLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
DB 1 MLPLHALLLLAAGAAARALEVPADGNAGLLAEPOIAMFCGKLNHMNVQNGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTVIQNWKCKGRKCKOCTHPHFVPIYRCVLG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTVIQNWKCKGRKCKOCTHPHFVPIYRCVLG 120

QY 121 EYVSDALLVPDKCKFLHQRMDVCTHLHWHVAVKETSCKSTNLHDYGMLLPGCIDKFR 180
DB 121 EYVSDALLVPDKCKFLHQRMDVCTHLHWHVAVKETSCKSTNLHDYGMLLPGCIDKFR 180

QY 181 GYVEFCCPLAESDNVDSADAEEDSDYVWGADTDYADGSEDKVVE---VAEEVEAEVE 238
DB 181 GYVEFCCPLAESDNVDSADAEEDSDYVWGADADYADGSDDKVVEEQDEDELTVE 240

QY 239 EEEADDDDEDGDDVVEEAEPEEATERTTSIATTTTTTSTESVEEYVREVCSEQAETG 298
DB 241 DEADADD-DEDDGDEI-BETEEYBEATERTTSIATTTTTTSTESVEEYVREVCSEQAETG 288

QY 299 PCRAMISRWDVTSKCAPFEYGGCGNRNFDTEECYMAVCGSAMQSLLKTTQEPPLA 358
DB 289 ----- 288

QY 359 RDPVKLPTTAASPDVADKYLETGPDENEHAHFQKAKERLEAKHREMSQVWREWEAEER 418

```



```

Db 289 -----VPTTAASDPDAVDKYLETPGDENEHAHFQKAKERLEAKHREMSQVWREWEAEER 343
QY 419 QAKNLPKADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYIT 478
Db 344 QAKNLPKADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYIT 403
QY 479 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSOVMTHLRVIY 538
Db 404 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSOVMTHLRVIY 463
QY 539 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLNMISEPRISYGNDAIMPST 598
Db 464 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLNMISEPRISYGNDAIMPST 523
QY 599 ETKTTVELLPVNGEFLDLPWHSEFGADSVDPANTENEVEPVDARPAADRGLTTRPGSL 658
Db 524 ETKTTVELLPVNGEFLDLPWHSEFGADSVDPANTENEVEPVDARPAADRGLTTRPGSL 583
QY 659 TNKTEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKGAIIIGLMVGGVVIATVIVI 718
Db 584 TNKTEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKGAIIIGLMVGGVVIATVIVI 643
QY 719 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQM 770
Db 644 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQM 695

RESULT 8
Q98SGO PRELIMINARY; PRT; 693 AA.
AC Q98SGO;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-amyloid precursor protein A.
GN APP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Van den Hurk W.H.;
RL Thesis (2001), Department of Biological Sciences,
RL University of Nijmegen, Nijmegen, Netherlands.
DR EMBL; AJ298150; CAC37193.1;
DR HSP; P05067; 1H23.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
FT SIGNAL.
SQ SEQUENCE 693 AA; 78568 MW; CAFIDF655CIAB653 CRC64;

Query Match 78.0%; Score 3163.5; DB 13; Length 693;
Best Local Similarity 79.1%; Pred. No. 1.1e-182;
Matches 611; Conservative 37; Mismatches 43; Indels 81; Gaps 5;

QY 1 MLPGLALLLAANWTARALEVPTDGNAGLLAEPOIAMFCGLNMHNMVQNGKWDSPSGTK 60
Db 1 MLPHTILLVITV-GALALEVPADNGGLLAEPQIAMFCGLNMHNMVQNGKWDSPSGTK 59
QY 61 TCIDTKEGILQYCOEYVPELQITNVNEANQPVTIQNCKRGRQCKTHPHFVPIYRCLVG 120
Db 60 GCIGTKEGILQYCOEYVPELQITNVNEANQPVTIQNCKRGRQCKSRTHVIVPYRCLVG 119
QY 121 EFVSALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180

```

```

Db 120 EFVSALLVPDKCKFLHQERMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 179
QY 181 GVEFFCCPLABESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEYA--EEBEVAEVE 238
Db 180 GVEFFCCPLABESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEYA--EEBEVAEVE 238
QY 239 EEEADDDDEDDGDEVEEAEPEEATERTSTIATTTTTTSESVEEVVREVCSEQAETG 298
Db 239 EEEADDDDEDDGDEVEEAEPEEATERTSTIATTTTTTSESVEEVVREVCSEQAETG 286
QY 239 PCRAMISWYFDVTGKCAPFFYGCCGGRNFDTEECVMAVCGSAMQSLLKTTQBPLA 358
Db 287 ----- 286
QY 359 RDPVKLPKTTAASDPDAVDKYLETPGDENEHAHFQKAKERLEAKHREMSQVWREWEAEER 418
Db 287 -----VPATAASTDPDAVDKYLETPGDENEHAHFQKAKERLEAKHREMSQVWREWEAEER 341
QY 419 QAKNLPKADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYIT 478
Db 342 QAKNLPKADKKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYIT 401
QY 479 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSOVMTHLRVIY 538
Db 402 ALQAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSOVMTHLRVIN 461
QY 539 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLNMISEPRISYGNDAIMPST 598
Db 462 ERMNQSLLYNPVAVEEIQDEVDLLOKEQNYSDVLNMISEPRISYGNDAIMPST 521
QY 599 ETKTTVELLPVNGEFLDLPWHSEFGADSVDPANTENEVEPVDARPAADRGLTTRPGSL 658
Db 522 ETKTTVELLPVNGEFLDLPWHSEFGADSVDPANTENEVEPVDARPAADRGLTTRPGSL 581
QY 659 TNKTEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKGAIIIGLMVGGVVIATVIVI 718
Db 582 TNKTEISEVKMDAEFRHDSGEYVHHQKLVFPAEDVGSNGKGAIIIGLMVGGVVIATVIVI 641
QY 719 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQM 770
Db 642 TLVMLKKKQYTSIHGGVVEVDAATPEERHLKSKMQONGYENPTYKFFEQM 693

RESULT 9
Q98SF9 PRELIMINARY; PRT; 695 AA.
AC Q98SF9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-amyloid precursor protein B.
GN APP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Van den Hurk W.H.;
RL Thesis (2001), Department of Biological Sciences,
RL University of Nijmegen, Nijmegen, Netherlands.
DR EMBL; AJ298151; CAC37194.1;
DR HSP; P05067; 1H23.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.

```


Tue Jan 28 14:22:52 2003

```
RESULT 11
O57394 PRELIMINARY; PRT; 699 AA.
ID O57394
AC O57394; 1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE EL amyloid precursor protein 699.
GN EL App699.
OS Narke japonica (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hymnosquala; Pristionotidae; Batoidae;
OC Torpediniformes; Narcinoidae; Narkidae; Narke.
OX NCBI_TaxID=62965;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ELECTRIC LOBE;
RX MEDLINE=98129705; PubMed=9461486;
RA Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
RA Suzuki T.;
RT "cDNA isolation of Alzheimer's amyloid precursor protein from
RT cholinergic nerve terminals of the electric organ of the electric
RT ray.";
RL Biochem. J. 330:29-33(1998).
DR EMBL; AB005544; BAA24230.1; -.
DR HSP; P05067; LH23.
DR InterPro; IPR001868; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF02177; A4_EXTRA; 1.
DR Pfam; PF03494; Beta-APP; 1.
DR PRINTS; PR00203; AMYLOIDA4.
DR SMART; SM00006; A4_EXTRA; 1.
DR PROSITE; PS00319; A4_EXTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
DR PROSITE; PS00320; A4_INTRA; 1.
SQ SEQUENCE 699 AA; 78879 MW; 952915C309D50E5C CRC64;

Query Match 71.8%; Score 2914; DB 13; Length 699;
Best Local Similarity 72.9%; Pred. No. 1.2e-167;
Matches 568; Conservative 59; Mismatches 58; Indels 94; Gaps 9;

QY 2 LPG-LALLLAAWTA-----RALEVPTDGNAGLL-AEPQIAMFCGRNLNMHNVQNGK 52
DB 5 LFGRLGMLLAAALVLAAPLCRALEVPTDGGAGLLAAEPQIAMFCGRNLNMHNVQNGK 64
QY 53 DSDPSGKTCTDKGILQYCEVYPELQITNVVEANQPTIOWCKRCKGCKTHPHFV 112
DB 65 VSDPSGTNTCFGTKEGILYCEVYVDPDQITNVVEANQPTIOWCKRCKGCKGHPHIV 124
QY 113 IPYRCVLGFEVDALLVPDKCKFLHQRMDVCETHLWHTVAKETCSKSTNLHDYGM 172
DB 125 VYRCVLGFEVDALLVPDKCKFLHQRMDVCETHLWHTVAKETCSKSTNLHDYGM 184
QY 173 PGCIDKFRGVFVCCPLAEESDNVDSADAEEDSDVWNGGADTDVADGSEDKVVEAE 232
DB 185 PGCIDFRGVFVCCPIPENDKIDS-DMDEEDSDVWNGGADTDVADGSEDKVVEAE 238
QY 233 EVAEVEEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 291
DB 239 KPIEEEEEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 292
QY 292 SEQAEETGPCRAMISRWFVDVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSQ 351
DB 293 ----- 292
QY 352 TTOEPLARDPVKLPTTAATDPADVLYETPGDENEHAHFQAKERLEAKHRERMSQV 411
DB 293 -----VPTTAATDPADVLYETPGDENEHAHFQAKERLEAKHRERMSQV 340
QY 412 EWEAEERQAKNLPKADKAVIQHFQKESLEQEAERQQLVETVHMARVEAMLRRL 471
DB 341 EWEAEERQAKNLPKADKAVIQHFQKESLEQEAERQQLVETVHMARVEAMLRRL 400
```


Query Match 69.8%; Score 2831; DB 4; Length 547;
Best Local Similarity 99.1%; Pred. No. 8.7e-163;
Matches 532; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60

QY 61 TCIDTKGIIQYQCEVYPELQITNVVEANOPVTIQNKCKRCKOCTHPHFVPIYRCLVG 120
DB 61 TCIDTKGIIQYQCEVYPELQITNVVEANOPVTIQNKCKRCKOCTHPHFVPIYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVAEVEE 240

QY 241 EADDEDEDEGEVEEAEPYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
DB 241 EADDEDEDEGEVEEAEPYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300

QY 301 RAMISRWFYVTEGCKAPFFYGGCGGNRNFDTEEYCMVCGSAMSLKTTQEPPLARD 360
DB 301 RAMISRWFYVTEGCKAPFFYGGCGGNRNFDTEEYCMVCGSAMSLKTTQEPPLARD 360

QY 361 PVKLPTTAASPPDAVDKYLETGPDENHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASPPDAVDKYLETGPDENHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420

QY 421 KNLPKADKAVIOHFQKESVLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 480
DB 421 KNLPKADKAVIOHFQKESVLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 480

QY 481 QAVPPRRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQTRSQVMTLRLVI 537
DB 481 QAVPPRRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQTRSQVMTLRLVI 537

DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.
KW Serine protease inhibitor.
SQ SEQUENCE 738 AA; 83577 MW; AF480F6D308FD298 CRC64;
Query Match 68.6%; Score 2785.5; DB 13; Length 738;
Best Local Similarity 69.9%; Pred. No. 7.1e-160;
Matches 540; Conservative 85; Mismatches 101; Indels 47; Gaps 14;

QY 5 LALLLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTKCID 64
DB 6 LALLLLAATAARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTKCID 65

QY 65 TREGILQYQCEVYPELQITNVVEANOPVTIQNKCKRCKOCTHPHFVPIYRCLVGFEVS 124
DB 66 NKEGILQYQCEVYPELQITNVVEANOPVTIQNKCKRCKOCTHPHFVPIYRCLVGFEVS 125

QY 125 DALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFRGVEF 184
DB 126 DALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSKSTNLHDYGMLLPCGIDKFRGVEF 185

QY 185 VCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVAEVEEADD 244
DB 186 VCCP-ADAKESAEEDSDVMWGGAEADYTENSMTR--DAAEAPVLE-DDEDADE 241

QY 245 DED-DEGD-----EVEEAEPYEATERTT-STATTTTTTSTESVEEVREVCSQAET 297
DB 242 EDEDDGDDGDRDEKTEEEEEE--EERTQSTSAALTSTTTTTSTESVEEVREVCSQAET 299

QY 298 GPCRAMISRWFYVTEGCKAPFFYGGCGGNRNFDTEEYCMVCGSAMSLKTTQEPPL 357
DB 300 GPCRAMISRWFYVTEGCKAPFFYGGCGGNRNFDTEEYCMVCGSAMSLKTTQEPPL 357

QY 358 ARDPVKLPPTAASPPDAVDKYLETGPDENHAHFQKAKERLEAKHRMSQVMREWEAE 417
DB 347 -----LPTSSPPDAVDKYLETGPDENHAHFQKAKERLEAKHRMSQVMREWEAE 400

QY 418 ROAKNLPKADKAVIOHFQKESVLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 477
DB 401 ROAKNLPKADKAVIOHFQKESVLEQEAANERQQLVETHMARVEAMLNDRRLALENITAL 460

QY 478 TALQAVPPRRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQTRSQVMTLRLVI 537
DB 461 SALQADPPRRHVNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQTRSQVMTLRLVI 520

QY 538 YERNQSLSLYVPAVAEIOEDVDELQEQNTYSDVLANNI SEPRISYCNDAIMP 597
DB 521 EERNQSLSLYVPAVAEIOEDVDELQEQNTYSDVLANNI SEPRISYCNDAIMP 579

QY 598 TETKTIVELLVNGEFLDQLPWHFSGADSVPAANTENEVEVPDARPAADRLTTPGSG 657
DB 580 T---ACLELLPAEDTQGFIFHP-ESFN---QPNTNQVEVPDARPAADRLTTPGSG 631

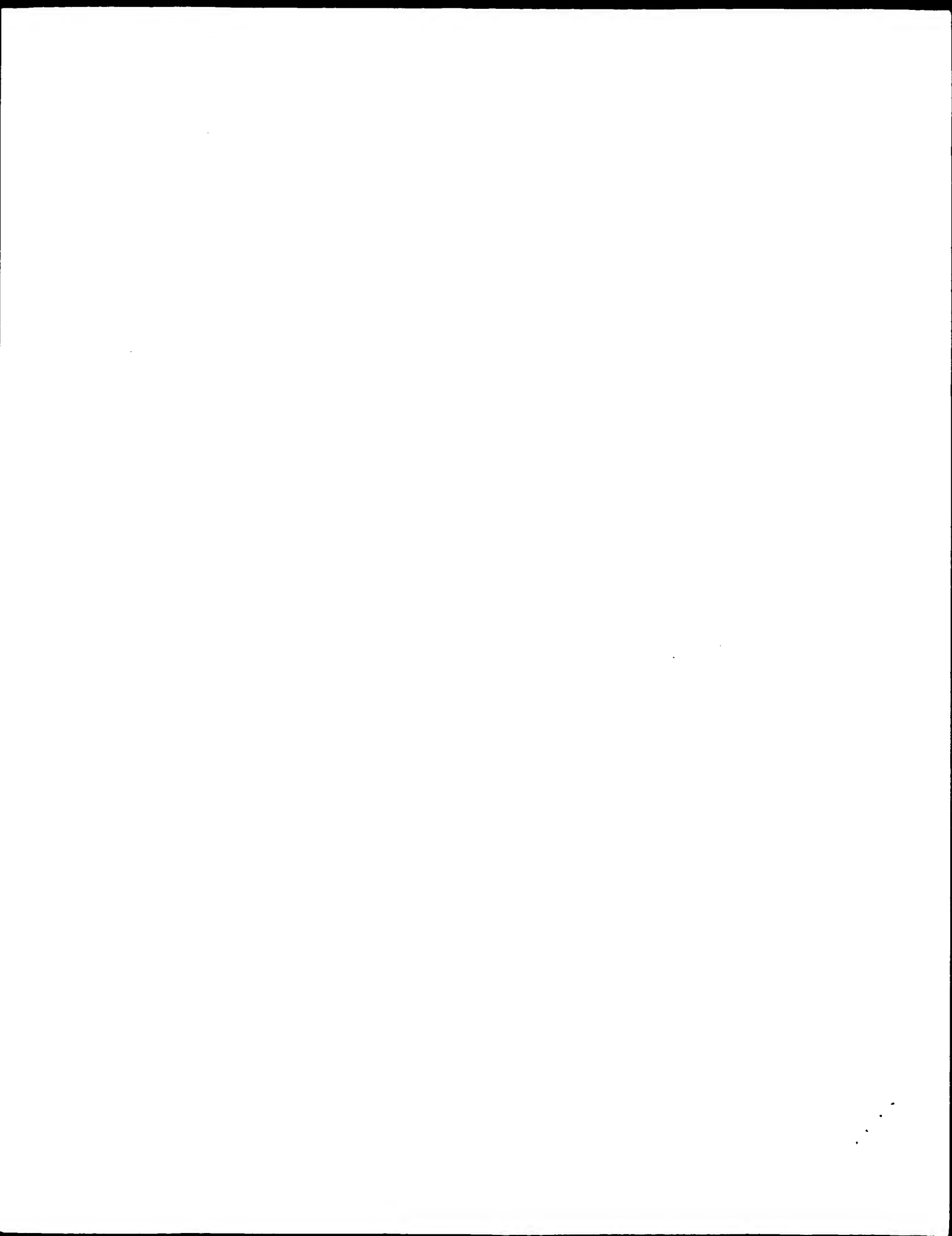
QY 658 LTNKITEETSEVKMDAEFRHDSGYEVHOKLVFFAEADVGSNGKAGIIGLMVGWVATVIV 717
DB 632 L---KPDDPELPELMEAEERHS---EYVHOKLVFFAEADVGSNGKAGIIGLMVGWVATVIV 685

QY 718 ITLVMLKKKQYTSIHGHVVEVDAAVTPERHLSKMOQNGYENPTYKFFSQMON 770
DB 686 ITLVMLKKKQYTSIHGHVVEVDAAVTPERHLSKMOQNGYENPTYKFFSQMON 738

RESULT 14
O73683
ID O73683 PRELIMINARY; PRT; 780 AA.
AC O73683;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
DE beta-amyloid protein (Beta-APP) (A-beta)].
GN APP.
OS Tetraodon fluviatilis (Puffer fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DR PROSITE; PS00320; A4_INTRA; 1.
FT NON_TER 1
SQ SEQUENCE 569 AA; 64753 MW; 0AB8B851863A19D CRC64;
Query Match 67.0%; Score 2717; DB 13; Length 569;
Best Local Similarity 82.5%; Pred. No. 6.8e-156;
Matches 534; Conservations 15; Mismatches 18; Indels 80; Gaps 5;
QY 126 ALLVPPKCKFLHQRMDVCEHLLHWHVAKETCSKSTNLHDYGMLLPGIDGKFRGVFV 185
Db 1 ALLVPPKCKLLHQRMDVCEHLLHWHVAKETCSKSMNLHDYGMLLSCGIDKFRGVFV 60
QY 186 CCPLAEESNVDSDADAEEDSDVWVGADTDYADGSEDKVVV--VAEEVEVAEVEEAD 243
Db 61 CCPLAEESNLDSDADAEEDSDVWVGADADYADGSDKVVVEEQPEDEBELIWEDEDAD 120
QY 244 DDEDDGDEVEEAEPEEYEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPCRAM 303
Db 121 DD-DDDDGDEI-BETEEVEEATERTTSIATTTTTTSTESVEEVV----- 163
QY 304 ISRWYFDVTGKCAPFYGGCGNRRNFDTEECYMAVCGSAMQSLSLLKTTQEP LARDPVK 363
Db 164 ----- 163
QY 364 LPTTAASTPDVADKYLETPGDENEHAHFOKAKERLEAKHRRMSQVWREWEAEARQAKNL 423
Db 164 VPTTAASTPDVADKYLETPGDENEHAHFOKAKERLEAKHRRMSQVWREWEAEARQAKNL 223
QY 424 PKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITALQAV 483
Db 224 PKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITALQTV 283
QY 484 PPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVPMVDPKAAQIRSQVMTHLRVYIERMNO 543
Db 284 PPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVPMVDPKAAQIRSQVMTHLRVYIERMNO 343
QY 544 SLSLLYNVPAVAEEIQDEVDDELQKEQNYSDVLANMISEPRISYGNALMPSLTETKT 603
Db 344 SLSFLYNVPAVAEEIQDEVDDELQKEQNYSDVLANMISEPRISYGNALMPTLTETKT 403
QY 604 VELLPVNGEFSLDLQFVHSGFADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTNKT 663
Db 404 VELLPVNGEFSLDLQFVHSGFADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTNKT 463
QY 664 EETSEVKMDAEFRHDSGYEVHOKLVFFAEVCGSNKGAIIGLMVGGVVATVIVITLVML 723
Db 464 EEVSEVKMDAEFRHDSGYEVHOKLVFFAEVCGSNKGAIIGLMVGGVVIANVIVITLVML 523
QY 724 KKKQYTSIHGVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMON 770
Db 524 KKKQYTSIHGVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMON 569

Search completed: January 28, 2003, 14:18:15
Job time : 42.9101 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 42.9063 seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLLAAWTARALEV.....KMQQNGYENPTYKFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	10 AAP94775	Novel amyloid prec
2	4058	100.0	770	14 AAR41546	Mutated APP770 exo
3	4058	100.0	770	15 AAR63442	Amyloid protein pr
4	4058	100.0	770	19 AAW40130	Human APP770 prote
5	4058	100.0	770	20 AAW97996	Human amyloid prec
6	4058	100.0	770	22 AAE10648	Human amyloid prot
7	4058	100.0	770	22 AAE11762	Human amyloid prec
8	4058	100.0	770	22 AAE06893	Human amyloid prec
9	4058	100.0	770	22 AAU06522	Human partial Amyl
10	4058	100.0	770	22 AAE02600	Human amyloid prec

11	4058	100.0	770	23 ABB78609	Human APP770 prote
12	4058	100.0	770	23 AAG58317	Human amyloid prec
13	4058	100.0	772	22 AAE10650	Human amyloid prot
14	4058	100.0	772	22 AAE06895	Human amyloid prec
15	4058	100.0	772	22 AAU06624	Human beta-amyloid
16	4058	100.0	772	22 AAU07223	Human amyloid prec
17	4058	100.0	772	22 AAE02602	Human amyloid prec
18	4058	100.0	772	23 ABB78611	Human APP770-KK pr
19	4054	99.9	770	15 AAR62505	Amyloid precursor
20	4053	99.9	770	18 AAW19500	APP770 mutant A-be
21	4053	99.9	770	18 AAW19497	APP770 mutant A-be
22	4053	99.9	770	18 AAW19485	APP770 mutant A-be
23	4053	99.9	770	18 AAW19482	APP770 mutant A-be
24	4053	99.9	770	22 AAE06913	Human amyloid prec
25	4050	99.8	770	22 AAE06912	Human amyloid prec
26	4047	99.7	768	23 AAU08959	Human amyloid beta
27	4047	99.7	770	18 AAW19506	APP770 mutant A-be
28	4047	99.7	770	18 AAW19491	APP770 mutant A-be
29	4046	99.7	770	13 AAR26340	APP770. Homo sapi
30	4046	99.7	770	18 AAW19488	APP770 mutant A-be
31	4046	99.7	770	18 AAW19503	APP770 mutant A-be
32	4039	99.5	770	11 AAR05717	NAP gene product a
33	3943.5	97.2	751	12 AAR10022	Beta-amyloid-relat
34	3943.5	97.2	751	13 AAR20328	Sequence encoded b
35	3943.5	97.2	751	20 AAY08615	Human beta-amyloid
36	3943.5	97.2	751	20 AAY08605	Human beta-amyloid
37	3943.5	97.2	751	22 AAE10649	Human amyloid prot
38	3943.5	97.2	751	22 AAE06894	Human amyloid prec
39	3943.5	97.2	751	22 AAU06623	Human partial Amyl
40	3943.5	97.2	751	22 AAE02601	Human amyloid prec
41	3943.5	97.2	751	23 AAO18050	Amyloid precursor
42	3943.5	97.2	751	23 ABB78610	Human APP751 prote
43	3943.5	97.2	751	23 AAG68316	Human amyloid prec
44	3943.5	97.2	753	22 AAE10651	Human amyloid prot
45	3943.5	97.2	753	22 AAE06896	Human amyloid prec

ALIGNMENTS

RESULT 1

AAP94775

ID AAP94775 standard; protein; 770 AA.

XX AC AAP94775;

XX DT 05-JUL-1990 (first entry)

XX DE Novel amyloid precursor protein (NAP).

XX KW SPAP; amyloid precursor protein; dysbolism; INS76; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Active-site 289...345

FT /*tag= a

FT /*label=INS76

XX PN EP304013-A.

XX PD 22-FEB-1989.

XX PF 16-AUG-1988; 88EP-0113283.

XX PR 15-AUG-1987; 87JP-0203298.

XX PR 21-AUG-1987; 87JP-0207995.

XX PR 18-NOV-1987; 87JP-0291404.

XX PR 11-DEC-1987; 87JP-0313228.

XX PR 05-FEB-1988; 88JP-0025260.

XX PR 10-FEB-1988; 88JP-0029366.

XX PR 19-FEB-1988; 88JP-0037905.

XX PR 25-MAY-1988; 88JP-0125660.


```

XX PA (ASAH ) ASahi KASEI KOGYO.
XX PI Kitaguchi N, Takahashi Y, Tokushima Y, Itoh H;
XX DR WPI; 1989-055458/08.
XX DR N-PSDB; AAN91049.
XX PT Human senile plaque amyloid precursor protein and DNA -
XX PT used for study and diagnosis of dysbolism in the central nervous
XX PT system.
XX PS Disclosure; ; 108pp; English.
XX CC AA. sequence from 289 to 345 encodes INS76 which is claimed as a senile
XX CC plaque amyloid precursor (SPAP), useful for diagnosis of dysbolism in the
XX CC CNS such as senile dementia.
XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMECGRLNMHMNVONGKWDSDPSGK 60
DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMECGRLNMHMNVONGKWDSDPSGK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVAEEVEE 240
QY 241 EADDEDEDGEVEEAEPEYBATERTTSIATTTTTTSTESVEEVREVCSEQAETGPC 300
DB 241 EADDEDEDGEVEEAEPEYBATERTTSIATTTTTTSTESVEEVREVCSEQAETGPC 300
QY 301 RAMLSRWYFDVTEGKCAPFFYGGCGGRNNDTEEYCMVCGSAMSGSLKTTQOEPLARD 360
DB 301 RAMLSRWYFDVTEGKCAPFFYGGCGGRNNDTEEYCMVCGSAMSGSLKTTQOEPLARD 360
QY 361 PVKLPPTTAASPDADVXYLETDPGDENEHAHFQAKERLEAKHRMSOVMREWEAEERQA 420
DB 361 PVKLPPTTAASPDADVXYLETDPGDENEHAHFQAKERLEAKHRMSOVMREWEAEERQA 420
QY 421 KNLPAKDKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
DB 421 KNLPAKDKAVIOHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
QY 481 QAVPPRRPHVFMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRIYER 540
DB 481 QAVPPRRPHVFMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVWTHLRIYER 540
QY 541 MNQSLSLYLVNPAVAEEIQDEVELLQKEQNSDDVLANNISPEPRISYGNLALMPSLTET 600
DB 541 MNQSLSLYLVNPAVAEEIQDEVELLQKEQNSDDVLANNISPEPRISYGNLALMPSLTET 600
QY 601 KTTVELLPVNGEFLSDLOLPWHSFGADSVPAANTENEVEPVDARPAADRGLITTRPGSGLN 660
DB 601 KTTVELLPVNGEFLSDLOLPWHSFGADSVPAANTENEVEPVDARPAADRGLITTRPGSGLN 660
QY 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNGKAIIGLMVGWGVVATVIVTL 720
DB 661 IKTEETSEVKMDAEFRHDSGYEVHHOKLVFFAEDVGSNGKAIIGLMVGWGVVATVIVTL 720
QY 721 VMLK$KQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770

```

```

DB 721 VMLK$KQYTSIHGGVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770
RESULT 2
AAR41546
ID AAR41546 standard; Protein; 770 AA.
XX AAR41546;
XX AC AAR41546;
XX DT 15-MAR-1994 (first entry)
XX DE Mutated APP770 exon 17 protein fragment.
XX KW Probe; mutation; exon 17; amyloid precursor protein; APP770;
XX KW substitution; progressive presenile dementia; Alzheimer's disease;
XX KW cerebral haemorrhage; cerebral amyloid angiopathy; 4 kD fragment;
XX KW blood vessels; brain parenchyma; assay; processing; plaque.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 692 /note= "Position of mutation which causes abnormal
FT processing of APP770"
XX PN EP561087-A.
XX PD 22-SEP-1993.
XX PF 20-MAR-1992; 92EP-0400771.
XX PR 20-MAR-1992; 92EP-0400771.
XX PA (INNO-) INNOGENETICS NV SA.
XX PI Cras P, Hendriks L, Martin J, Van Broeckhoven C;
XX DR WPI; 1993-296442/38.
XX DR N-PSDB; AAQ48860.
XX PT New mutant form of beta-amyloid polypeptide - related to
XX PT development of cerebral haemorrhage and Alzheimer's disease, also
XX PS corresp. nucleic acid, vectors, host cells and antibodies
XX PS Disclosure; Fig 4; 21pp; English.
XX CC This sequence is encoded by exon 17 of the amyloid precursor protein
XX CC APP770. A mutation at position 2075 of the DNA sequence, a C>G
XX CC substitution, causes the substitution of Ala for Gly at codon 692.
XX CC This mutation has been detected in related patients with progressive
XX CC presenile dementia (Alzheimer's disease) or cerebral haemorrhage due
XX CC to cerebral amyloid angiopathy. It may be responsible for the
XX CC deposition of a 4 kD proteolytic fragment of APP in blood vessel
XX CC walls and brain parenchyma. Probes specific for the mutation (see
XX CC also AAQ48858-59) can be used to assay mRNA encoding substances which
XX CC cause abnormal processing of APP related to plaque formation, and to
XX CC detect this specific mutation.
XX SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 14; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMECGRLNMHMNVONGKWDSDPSGK 60
DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPQIAMECGRLNMHMNVONGKWDSDPSGK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIQNCKRGRKCKTHPHFVPIYRCLVG 120

```


Db 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVPDARPAADRLTTRPGSLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
 QY 721 VMLKKQYTSIHGGVVEVDAAVTPERHLSKMQQNGYENPTYKFFEQMN 770
 Db 721 VMLKKQYTSIHGGVVEVDAAVTPERHLSKMQQNGYENPTYKFFEQMN 770

RESULT 4
 AAW40130 standard; Protein: 770 AA.
 XX
 AC AAW40130;
 XX
 DT 03-JUN-1998 (first entry)
 XX
 DE Human APP770 protein.
 XX
 KW Amyloid-beta peptide; A-beta; beta-AP; senile plaque; angiopathy;
 KW brain; membrane-spanning glycoprotein; beta-amyloid precursor protein;
 KW APP770; chromosome 21; human; Alzheimers disease; AD; amyloid filament;
 KW treatment; disease; Down's syndrome; hereditary cerebral haemorrhage.
 XX
 OS Homo sapiens.
 XX
 PN WO9748983-A1.
 XX
 PD 24-DEC-1997.
 XX
 XX 18-JUN-1997; 97WO-US10601.
 XX
 XX 18-JUN-1996; 96US-0665649.
 XX
 PA (ATHE-) ATHENA NEUROSCIENCES INC.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 XX
 PI Citron M, Schenk D, Selkoe DJ, Seubert PA;
 XX
 DR WPI: 1998-063287/06.
 DR N-PSDB; AAV10322.
 XX
 PT Identifying compounds that alter cellular production of amyloid-beta
 PT 42 fragment - in vitro or in transgenic animal models, potentially
 PT useful for treatment of Alzheimer's and other amyloid deposition
 PT diseases
 XX
 PS Disclosure; Fig 10; 86pp; English.
 XX

CC This sequence represents the human beta-amyloid precursor protein
 CC APP770, which is a membrane-spanning glycoprotein encoded by a gene
 CC on the long arm of chromosome 21. A fragment of the APP protein is known
 CC as the amyloid-beta peptide (A-beta), also known as the beta-Ap peptide,
 CC which forms the subunit of the amyloid filaments comprising senile
 CC (amyloid) plaques and the amyloid deposits in small cerebral and
 CC meningeal blood vessels (amyloid angiopathy). The A-beta peptide can be
 CC a 39-43 amino acid fragment. This invention provides methods of screening
 CC compounds for their ability to alter the production of the A-beta
 CC peptide, which is composed of >41 amino acids, alone, or in combination
 CC with the A-beta peptide composed of 40 amino acids or less. Such agents
 CC that reduce the production of the A-beta peptide are potentially useful
 CC for treatment of Alzheimers Disease or other diseases involving amyloid
 CC deposition such as Down's syndrome, hereditary cerebral haemorrhage with
 CC amyloidosis of Dutch type and advanced aging of the brain.
 XX
 XX Sequence 770 AA;
 Query Match 100.0%; Score 4058; DB 19; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5,3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSDPGTK 60
 Db 1 MLPGLALLLLAANTARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNQVNGKWDSDPGTK 60
 QY 61 TCIDTKEGILQYCEVYPELOITNVVEANQPVITONCKRCKOCKTHPHFVIPYRCLVG 120
 Db 61 TCIDTKEGILQYCEVYPELOITNVVEANQPVITONCKRCKOCKTHPHFVIPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPGCIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPGCIDKFR 180
 QY 181 GVERVCCPLABESDNVDSADAEEDSDVWVGADTDYADGEDKVVEAEVEEVEE 240
 Db 181 GVERVCCPLABESDNVDSADAEEDSDVWVGADTDYADGEDKVVEAEVEEVEE 240
 QY 241 EADDDDEDDGDEVEEAEPEYEEATETTTTATTTTTTETSEVVEVREVCSEAEATGPC 300
 Db 241 EADDDDEDDGDEVEEAEPEYEEATETTTTATTTTTTETSEVVEVREVCSEAEATGPC 300
 QY 301 RAMISRWFYDTEGKCAPFFYGGCGGNRNNDTEYCMVCGSAMSQSLKTTQEPPLARD 360
 Db 301 RAMISRWFYDTEGKCAPFFYGGCGGNRNNDTEYCMVCGSAMSQSLKTTQEPPLARD 360
 QY 361 PVKLPPTTAASTPDADV KYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
 Db 361 PVKLPPTTAASTPDADV KYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
 QY 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMNDRRLALENYITAL 480
 Db 421 KNLPKADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMNDRRLALENYITAL 480
 QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 540
 Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRWVDPKAAQIRSQVMTHLRVIER 540
 QY 541 MNQSLSLLYNPVPAVEETQDEVDLQKEQYSDDLANMISEPRISYNDALMPSLTET 600
 Db 541 MNQSLSLLYNPVPAVEETQDEVDLQKEQYSDDLANMISEPRISYNDALMPSLTET 600
 QY 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVPDARPAADRLTTRPGSLTN 660
 Db 601 KTTVELLPVNGEESLDDLPWHSFGADSVDPANTENEVPDARPAADRLTTRPGSLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
 QY 721 VMLKKQYTSIHGGVVEVDAAVTPERHLSKMQQNGYENPTYKFFEQMN 770
 Db 721 VMLKKQYTSIHGGVVEVDAAVTPERHLSKMQQNGYENPTYKFFEQMN 770

RESULT 5
 AAW97996 standard; Protein: 770 AA.
 ID AAW97996
 AC AAW97996;
 XX
 XX 21-JUN-1999 (first entry)
 DT
 XX Human amyloid precursor protein.
 DE
 XX Amyloid precursor protein; App; human; gene targeting;
 KW homologous recombination; transgenic mouse; transgenic animal;
 KW animal model; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 672..711
 FT /note="beta-amyloid domain"
 XX

PN WO9909150-A1.
 XX 25-FEB-1999.
 XX 18-AUG-1997; 97WO-US14507.
 XX 18-AUG-1997; 97WO-US14507.
 XX (FARB) BAYER CORP.
 XX Wirak DO;
 XX WPI; 1999-181029/15.
 DR Modification of target nucleic acids - by homologous recombination,
 XX particularly for introducing a humanised amyloid precursor
 XX protein gene into rodents for producing models of Alzheimer's
 XX disease
 XX Disclosure; Page 85-88; 209pp; English.
 XX This polypeptide comprises human amyloid precursor protein (hAPP).
 XX The invention provides a novel gene targeting strategy that
 XX facilitates the introduction of one or more specific mutations
 XX into any gene in a single double reciprocal homologous recombination
 XX step. The method has been used particularly for introducing a
 XX humanised APP gene into rodents for producing animal models of
 XX Alzheimer's disease (AD). 4 Independent lines of transgenic mice
 XX (lines ES5007, ES5103, ES5401 and ES5403) have been created using
 XX the novel gene targeting technique applied to embryonic stem cells.
 XX In each line, the mouse APP gene was modified to encode a
 XX mouse/human hybrid (m/hAPP) where amino acid residues 666-770 of
 XX APP770 are encoded by human cDNA sequences instead of mouse genomic
 XX exons (exons 16-18). Within these residues, only 3 amino acid
 XX differences exist between the mouse and human proteins, i.e.
 XX Gly-676 to Arg, Phe-681 to Thr and Arg-684 to His. This exon-cDNA
 XX fusion gene therefore encodes an APP containing a humanised
 XX beta-amyloid domain. Swedish, London, Swedish/London and stop
 XX mutations have also been introduced. Targetting vector sequences
 XX are provided (see AAX24730-33).
 XX Sequence 770 AA;
 S0

Query Match 100.0%; Score 4058; DB 20; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMFCGLRNHMMNVQNGKWDSPSGTK 60
 DB 1 MLPGLALLLLAATARALEVPTDGNAGLLAEPQIAMFCGLRNHMMNVQNGKWDSPSGTK 60
 QY 61 TCIDTREGILQYCOEYVPELQITNVVEANOPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 DB 61 TCIDTREGILQYCOEYVPELQITNVVEANOPVTIQNCKRGRKCKTHPHFVTPYRCLVG 120
 QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSKSTNLHDYGMLLPCGIDKFR 180
 DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETCSKSTNLHDYGMLLPCGIDKFR 180
 QY 181 GVFEVCCPLAEESDNVDSADAEDSDSVMMGGADTDYADGSEDKVVEAEVEAEVEE 240
 DB 181 GVFEVCCPLAEESDNVDSADAEDSDSVMMGGADTDYADGSEDKVVEAEVEAEVEE 240
 QY 241 EADDEDEDEGDEVEEAEPYEATERTSIATTTTTTSTVESVEEVVRECVCSQATGTC 300
 DB 241 EADDEDEDEGDEVEEAEPYEATERTSIATTTTTTSTVESVEEVVRECVCSQATGTC 300
 QY 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTQEP LARD 360
 DB 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTQEP LARD 360
 QY 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQAKERLEAKHRMSQVNRWEAEERQA 420
 XXX

Db 361 PVKLPPTAASTPDVADKYLETPGDENEHAHFQAKERLEAKHRMSQVNRWEAEERQA 420
 QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENY ITAL 480
 Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENY ITAL 480
 QY 481 QAVPPRPRIHVFNMLKKYVRAEQDKRQHTLKHFEHVPMVDPKKAQIRSQVMTHLRVIYER 540
 Db 481 QAVPPRPRIHVFNMLKKYVRAEQDKRQHTLKHFEHVPMVDPKKAQIRSQVMTHLRVIYER 540
 QY 541 MNQSLSLYNVPAVAEEIODEVDELQKQNYSDOVLANNISEPRIISYNDALMPSLTET 600
 Db 541 MNQSLSLYNVPAVAEEIODEVDELQKQNYSDOVLANNISEPRIISYNDALMPSLTET 600
 QY 601 KTTVELLPVNGEFSDDLQPHWSFGADSVANTENEVEPVDARPAADRLGTTTPGSGGLTN 660
 Db 601 KTTVELLPVNGEFSDDLQPHWSFGADSVANTENEVEPVDARPAADRLGTTTPGSGGLTN 660
 QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEEDVGSNKGAIIGLMVGGVVIATVITVL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEEDVGSNKGAIIGLMVGGVVIATVITVL 720
 QY 721 VMLKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 Db 721 VMLKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
 RESULT 6
 AAE10648
 ID AAE10648 standard; Protein; 770 AA.
 XX AAE10648;
 AC AAE10648;
 XX 10-DEC-2001 (first entry)
 DT Human amyloid protein precursor 770 (APP770) isoform.
 DE Human; aspartyl protease 1; Aspl; amyloid precursor protein 770; APP770;
 KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
 KW amyloid plaque; neuronal loss; proteolytic; nontropic; neuroprotective.
 XX Homo sapiens.
 OS GB2357767-A.
 PN 04-JUL-2001.
 XX 22-SEP-2000; 2000GB-0023315.
 PF 23-SEP-1999; 99US-0155493.
 PR 23-SEP-1999; 99US-0404133.
 PR 23-SEP-1999; 99WO-US20881.
 PR 13-OCT-1999; 99US-0416901.
 PR 06-DEC-1999; 99US-0169232.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA Bienkowski MJ, Gurney M;
 PI WPI; 2001-444208/48.
 DR N-PSDB; AAD17897.
 XX Polypeptide comprising fragments of human aspartyl protease with
 PT amyloid precursor protein processing activity and alpha-secretase
 PT activity, for identifying modulators useful in treating Alzheimer's
 PT disease -
 XX Example 8; Page 142-144; 187pp; English.
 PS The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
 CC Aspl proteins which lack transmembrane domain or amino terminal
 CC domain or cytoplasmic domain and retains alpha-secretase activity
 CC and amyloid protein precursor (APP) processing activity. The proteins

CC of the invention are useful for assaying hu-Asp1 alpha-secretase
CC activity, which in turn is useful for identifying modulators of
CC hu-Asp1 alpha-secretase activity, where modulators that increase
CC hu-Asp1 alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC neuronal loss. Hu-Asp1 protease substrate is useful for assaying
CC hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Asp1 proteolytic activity. The present sequence is amyloid protein
CC precursor 770 (APP770) isoform from human.
xx

Qy	1	MLPGLALLLLAATARA	LEVPTDGNAGLLAEPQI	AMFCGR	LNHMNVQNGK	WSDSPSGTK	60
Db	1	MLPGLALLLLAATARA	LEVPTDGNAGLLAEPQI	AMFCGR	LNHMNVQNGK	WSDSPSGTK	60
Qy	61	TCIDTKEGILQYCEV	PELQITNVVANQPV	TIQNKCKR	GKCKTHPHFV	IPYRCLVG	120
Db	61	TCIDTKEGILQYCEV	PELQITNVVANQPV	TIQNKCKR	GKCKTHPHFV	IPYRCLVG	120
Qy	121	EFVSDALLVPDKCF	LHQRMDVCETHLH	WHTVAKET	CEKSTNLHDY	GMLPGIDKFR	180
Db	121	EFVSDALLVPDKCF	LHQRMDVCETHLH	WHTVAKET	CEKSTNLHDY	GMLPGIDKFR	180
Qy	181	GVEFYCCPLAESD	NVDSADAEEDSD	VMWGGADTD	ADGSEDKV	VEAEVEE	240
Db	181	GVEFYCCPLAESD	NVDSADAEEDSD	VMWGGADTD	ADGSEDKV	VEAEVEE	240
Qy	241	EADDEDEDDGVE	EEAEPEYEA	TERTTSIAT	TTTTT	TESVEEVREVC	300
Db	241	EADDEDEDDGVE	EEAEPEYEA	TERTTSIAT	TTTTT	TESVEEVREVC	300
Qy	301	RAMISRWFVDT	TEGKCAPFFY	GCGGNRNF	TEYCMV	CGSAMSOSLLK	360
Db	301	RAMISRWFVDT	TEGKCAPFFY	GCGGNRNF	TEYCMV	CGSAMSOSLLK	360
Qy	361	PVKLTPTAAS	TPDAVDKYLET	PGDENEHA	HFQAKERLE	AKHRMSQVMRE	420
Db	361	PVKLTPTAAS	TPDAVDKYLET	PGDENEHA	HFQAKERLE	AKHRMSQVMRE	420
Qy	421	KNLPKADKKAV	IQHFQEKVES	LEQEAANE	ROQLVETH	MAVEMLNDRRL	480
Db	421	KNLPKADKKAV	IQHFQEKVES	LEQEAANE	ROQLVETH	MAVEMLNDRRL	480
Qy	481	QAVPPRPRH	VFNLMKKY	RAEOKDRH	TLKHFEH	VRMVDPKAAQ	540
Db	481	QAVPPRPRH	VFNLMKKY	RAEOKDRH	TLKHFEH	VRMVDPKAAQ	540
Qy	541	MNQSLSLN	YNPAVAEE	IQDEVEL	LOKEQY	SDDVLANMI	600
Db	541	MNQSLSLN	YNPAVAEE	IQDEVEL	LOKEQY	SDDVLANMI	600
Qy	601	KTTVELLPNG	EFLDDLP	WHFSFG	ADSV	PANTEVEP	660
Db	601	KTTVELLPNG	EFLDDLP	WHFSFG	ADSV	PANTEVEP	660
Qy	661	IKTEISEVK	MDAEPH	DSGVEV	HQKLV	FFAEADV	720
Db	661	IKTEISEVK	MDAEPH	DSGVEV	HQKLV	FFAEADV	720
Qy	721	VMLKKQY	TSIHGGV	VEVDA	AVTP	PERHLSK	770
Db	721	VMLKKQY	TSIHGGV	VEVDA	AVTP	PERHLSK	770

RESULT 7
AAE11762

ID	AAE11762 standard; Protein; 770 AA.
XX	AAE11762;
AC	18-DEC-2001 (first entry)
XX	Human amyloid precursor protein (APP).
DE	Human; neuroprotective; nootropic; immunostimulant; Alzheimer's disease;
XX	anticonvulsant; vaccine; gene therapy; Pick's disease; antidiabetic;
KW	systemic amyloidosis; maturity onset diabetes; Parkinson's disease;
KW	Huntington's disease; fronto-temporal dementia; encephalopathy; ALS;
KW	amyotrophic lateral sclerosis; amyloid precursor protein; APP.
XX	Homo sapiens.
OS	
XX	Key
XX	Location/Qualifiers
FT	Peptide
FT	1..18
FT	/label= Signal_peptide
FT	19..770
FT	/label= Mature_human_AAP_protein
FT	18..700
FT	/label= Extracellular_domain
FT	672..714
FT	/note= "Abeta-42/43 core peptide"
FT	700..723
FT	/label= Transmembrane_domain
FT	714..770
FT	/note= "C-100 fragment"
FT	723..770
FT	/label= Intracellular_domain
XX	WO200162284-A2.
PN	
XX	30-AUG-2001.
PD	
XX	19-FEB-2001; 2001WO-DK00113.
PF	
XX	21-FEB-2000; 2000DK-0000265.
PR	
XX	01-MAR-2000; 2000US-186295P.
XX	(NEBI-) M & E BIOTECH AS.
PA	
XX	Birk P, Jensen MR, Nielsen KG;
PI	
XX	WPI; 2001-589796/66.
DR	
XX	N-PSDB; AAD18754.
DR	
XX	In vivo down-regulation of amyloid protein for the treatment of
PT	Alzheimer's, comprises presenting an amyloidogenic polypeptide or its
PT	subsequence and/or at least one analogue of the amyloidogenic
PT	polypeptide to the immune system -
XX	
PS	Claim 23; Page 113-116; 120pp; English.
XX	
XX	The invention relates to a method for in vivo down-regulation of amyloid
CC	protein such as beta amyloid (Abeta) in an animal, including human. The
CC	method comprising presenting to the animal's immune system an
CC	immunogenically effective amount of at least one amyloidogenic protein
CC	or its subsequence and/or at least one analogue of the amyloidogenic
CC	polypeptide. The amyloidogenic protein or its subsequence, and its
CC	analogue is useful for the preparation of an immunogenic composition
CC	comprising an adjuvant for down-regulating amyloid in an animal. They are
CC	also useful in the treatment, prophylaxis or amelioration of Alzheimer's
CC	disease or other diseases characterised by amyloid deposits. They are
CC	also useful in the treatment of systemic amyloidosis, maturity onset
CC	diabetes, Parkinson's disease, Huntington's disease, fronto-temporal
CC	dementia, amyotrophic lateral sclerosis (ALS), Pick's disease and
CC	prior-related transmissible spongiform encephalopathies. They are also
CC	useful for inducing production of antibodies against an amyloidogenic
CC	polypeptide. The present sequence is human amyloid precursor protein
CC	(APP).
XX	

SQ Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSDPSGK 60
 Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Qy 61 TCIDTKEGILQYCOEVPPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120
 Db 61 TCIDTKEGILQYCOEVPPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHOERMDVCETHLHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
 Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
 Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240

Qy 241 EADDEDEDEGDEVEEAEPYEAERTTSIATTTTTTESVEEVVREVCSEAETGPC 300
 Db 241 EADDEDEDEGDEVEEAEPYEAERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

Qy 301 RAMLSRWYFDVTGKCAPFYGGCGGNRNNFDTEECYMAVCGSAMQSLLKTTOEPLARD 360
 Db 301 RAMLSRWYFDVTGKCAPFYGGCGGNRNNFDTEECYMAVCGSAMQSLLKTTOEPLARD 360

Qy 361 PVKLPTTAATPDVADVKILETPGDENEHAFQKAKERLEAKHRMSQVREWEAEROA 420
 Db 361 PVKLPTTAATPDVADVKILETPGDENEHAFQKAKERLEAKHRMSQVREWEAEROA 420

Qy 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 Db 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

Qy 481 QAVPPRPHRVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSOVMTHLRVIYER 540
 Db 481 QAVPPRPHRVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSOVMTHLRVIYER 540

Qy 541 MNGSLLYNVPAAVEIQDEVDLQKQNYSDVLANMISEPRISYNDALMPSLTET 600
 Db 541 MNGSLLYNVPAAVEIQDEVDLQKQNYSDVLANMISEPRISYNDALMPSLTET 600

Qy 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEVPDARPAADRLGTLTPGSGLTN 660
 Db 601 KTTVELLPVNGEFLDLPWHSFGADSVPAANTEVEVPDARPAADRLGTLTPGSGLTN 660

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHOKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
 Db 661 IKTEEISEVKMDAEFRHDSGYEVHHOKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720

Qy 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMON 770
 Db 721 VMLKKQYTSIHGGVVEVDAVTPPERHLSKMQQNGYENPTYKFFEQMON 770

RESULT 8

AAE06893

ID AAE06893 standard; Protein: 770 AA.

XX AC AAE06893;

XX DT AAE06893;

XX 23-OCT-2001 (first entry)

DE Human amyloid precursor protein 770 (APP770) isoform.

XX

KW Human; aspartyl protease; beta-amyloid precursor protein 770; APP770;
 KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
 KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; neurotropic;
 KW neuroprotective; antisense therapy; gene therapy; chromosome 21.

XX

OS Homo sapiens.

XX WO200150829-A2.

XX 19-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00799.

XX 09-MAY-2001; 2001WO-IB00799.

XX (BIEN/) BIENKOWSKI M J.

XX (GURN/) GURNEY M E.

XX (HEIN/) HEINRIKSON R L.

XX (PARO/) PARODI L A.

XX (YANR/) YAN R.

XX

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX WPI; 2001-483072/52.

XX DR N-PSDB; AAD13278.

XX

PT Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity -

PS Claim 8; Page 171-173; 185pp; English.

XX

CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
 CC precursor protein (APP) isoforms and their corresponding DNA molecules.
 CC Human aspartyl proteases can act as beta-secretase proteases useful for
 CC treating Alzheimer's disease. APP isoforms are useful for identifying
 CC modulators of amyloid-beta peptide production, for use in designing
 CC therapeutics for the treatment and prevention of Alzheimer's disease,
 CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
 CC and neuronal loss. APP isoforms are also used in methods for identifying
 CC inhibitors and modulators of human Asp2 activity. The invention relates
 CC to a method for identifying agents that modulate the activity of human
 CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
 CC as a means to screen in cellular assays for the inhibitors of beta- and
 CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
 CC polymerase chain reactions (PCR). The probes are useful for detecting
 CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
 CC blots. The present sequence is human wild-type amyloid precursor
 CC protein 770 (APP770) isoform. APP770 gene is localised of chromosome 21.

XX

SQ Sequence 770 AA;

Query Match

Best Local Similarity 100.0%; Score 4058; DB 22; Length 770;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Db 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRLLNMHNVQNGKWDSDPSGK 60

Qy 61 TCIDTKEGILQYCOEVPPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCOEVPPELQITNVVEANQPVTIONCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCKFLHOERMDVCETHLHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240

Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240

Qy 241 EADDEDEDEGDEVEEAEPYEAERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

Db 241 EADDEDEDEGDEVEEAEPYEAERTTSIATTTTTTESVEEVVREVCSEAETGPC 300

QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQPEPLARD 360
 DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQPEPLARD 360
 QY 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVNRWEAEERQA 420
 DB 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVNRWEAEERQA 420
 QY 421 KNLPAKDKKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 DB 421 KNLPAKDKKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
 QY 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRIYIER 540
 DB 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRIYIER 540
 QY 541 MNQSLSLLYNVPVAEIEQDEVELLOKEQNYSDVLANNMISEPRISYGNDAIMPSTET 600
 DB 541 MNQSLSLLYNVPVAEIEQDEVELLOKEQNYSDVLANNMISEPRISYGNDAIMPSTET 600
 QY 601 KTTVELLPVNGEFSLDLQPNHSGADSVPAENTENEVEPVDARPAADRGTLTRPGSLTN 660
 DB 601 KTTVELLPVNGEFSLDLQPNHSGADSVPAENTENEVEPVDARPAADRGTLTRPGSLTN 660
 QY 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSGNKGAIIGLMVGGVVIATVIVITL 720
 DB 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSGNKGAIIGLMVGGVVIATVIVITL 720
 QY 721 VMLKKKQVTSIHGKVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
 DB 721 VMLKKKQVTSIHGKVEVDAAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 9
 AAU06622
 ID AAU06622 standard; Protein; 770 AA.

XX AAU06622;

DT 24-OCT-2001 (first entry)

XX Human partial Amyloid precursor protein, APP770.

KW Human; Aspartyl protease; Asp2; beta-secretase; neurotropic;
 KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta; APP770.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 358

ET /note= "Encoded by GCC"

XX WO200149098-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

DR WPI; 2001-502549/55.

DR N-PSDB; AAS11549.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity -
 XX Disclosure; Page 171-173; 185pp; English.
 XX The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp
 CC proteins and vectors expressing them, and a polypeptide (isoform of
 CC amyloid protein precursor (APP)) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
 CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease.
 CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
 CC is associated with increased levels of Abeta processing is useful in
 CC assays relating the Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridize to
 CC Asp oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is human APP770.

XX Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPIQIAFCGRLLNMHNQVNGKWDSPSGTK 60

DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPIQIAFCGRLLNMHNQVNGKWDSPSGTK 60

QY 61 TCIDTRKREGILOQCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCVLVG 120

DB 61 TCIDTRKREGILOQCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCVLVG 120

QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180

DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180

QY 181 GVEFVCCPLAESDNVDSDADEDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240

DB 181 GVEFVCCPLAESDNVDSDADEDDSDVWVGADTDYADGSEDKVVEVAEEVEEVEE 240

QY 241 EADDDDEDDGDEVEEAPEEATERTSTATTTTTTSTESVEEVVRCVSEQAETGPC 300

DB 241 EADDDDEDDGDEVEEAPEEATERTSTATTTTTTSTESVEEVVRCVSEQAETGPC 300

QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQPEPLARD 360

DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQPEPLARD 360

QY 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVNRWEAEERQA 420

DB 361 PVKLPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVNRWEAEERQA 420

QY 421 KNLPAKDKKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

DB 421 KNLPAKDKKAVIOHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480

QY 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRIYIER 540

DB 481 QAVPPRRHVFNMMLKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRIYIER 540

QY 541 MNQSLSLLYNVPVAEIEQDEVELLOKEQNYSDVLANNMISEPRISYGNDAIMPSTET 600


```

Db 541 MNQSLSLLYNPVAAVEEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
QY 601 KTTVELLPVNGEFLSDDLQPHWSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLSDDLQPHWSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIIGLMVGGVVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIIGLMVGGVVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 10
AAE02600
ID AAE02600 standard; Protein: 770 AA.
XX
AC AAE02600;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human amyloid precursor protein 770 (APP 770) protein.
XX
KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW Alzheimer's disease; antialzheimer's.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 358 /note= "Encoded by GGC"
FT
XX
XX WO200123533-A2.
XX
XX 05-APR-2001.
XX
XX 22-SEP-2000; 2000WO-US26080.
XX
XX 23-SEP-1999; 99US-0155493.
XX 23-SEP-1999; 99WO-US20881.
XX 13-OCT-1999; 99US-0416901.
XX 06-DEC-1999; 99US-0169232.
XX
XX (PHAA ) PHARMACIA & UPJOHN CO.
XX
XX Gurney M, Bienkowski MJ;
XX
XX WPI; 2001-290516/30.
XX N-PSDB; AAD06770.
XX
XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
XX protein, useful for the treatment of Alzheimer's disease -
XX
XX Example 8; Page 170-172; 189pp; English.
XX
XX The present invention relates to enzymes for cleaving the alpha-
XX secretase site of the amyloid precursor protein (APP) and methods of
XX identifying those enzymes. The methods may be used to identify enzymes
XX that may be used to cleave the alpha-secretase cleavage site of the APP
XX protein. The enzymes may be used to treat or modulate the progress of
XX Alzheimer's disease. The present sequence is human APP 770 protein.
XX
XX Sequence 770 AA;
XX
XX Query Match 100.0%; Score 4058; DB 22; Length 770;
XX Best Local Similarity 100.0%; Pred. No. 5.3e-289;
XX Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MLPGLALLLAWTARALEVPTDGNAGLLAEFQIAMFCGRLNMHNVQNGKWDSPSGTK 60
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

```

Db 1 MLPGLALLLAWTARALEVPTDGNAGLLAEFQIAMFCGRLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCOEYVPELQITNNVVEANQPVTIQNWCKRGRKCKOCTHFFHVPYRCLVG 120
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 61 TCIDTREGILQYCOEYVPELQITNNVVEANQPVTIQNWCKRGRKCKOCTHFFHVPYRCLVG 120
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 121 EFVSDALLVPDKCFHJQERMDVCEPHLHWHHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 121 EFVSDALLVPDKCFHJQERMDVCEPHLHWHHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGADTDYADGSEDKVVEVAEEVEAEVEE 240
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVMWGADTDYADGSEDKVVEVAEEVEAEVEE 240
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 301 RAMISRWYFDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMQSLSLLKTTQBPPLARD 360
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 301 RAMISRWYFDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMQSLSLLKTTQBPPLARD 360
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 361 PVKLPPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 361 PVKLPPTTAASDPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 421 KNLKADKAKVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 421 KNLKADKAKVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 481 QAVPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVMVDPKKAAQIRSOVMTHLRVIYER 540
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 481 QAVPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVMVDPKKAAQIRSOVMTHLRVIYER 540
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 541 MNQSLSLLYNPVAAVEEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 541 MNQSLSLLYNPVAAVEEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 601 KTTVELLPVNGEFLSDDLQPHWSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 601 KTTVELLPVNGEFLSDDLQPHWSFGADSVSPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIIGLMVGGVVIATVITL 720
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIIGLMVGGVVIATVITL 720
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

RESULT 11
ABB78609
ID ABB78609 standard; Protein: 770 AA.
XX
AC ABB78609;
XX
XX 16-JUL-2002 (first entry)
XX
DE Human APP770 protein sequence SEQ ID NO:55.
XX
KW Human; Asp-1; Asp-2; aspartyl protease; Alzheimer's disease;
KW proteolytic; amyloid precursor protein; APP.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 358 /note= "encoded by GGC"
XX FT
XX
XX GB2367060-A.
XX
XX 27-MAR-2002.
XX
PD

```


thè production of beta-amyloid by contacting with beta-amyloid producing cells; and (2) screening kits. (1) have neurotropic and neuroprotective activities. (1) suppress the phosphorylation of amyloid precursor protein (APP) which is an essential step in the production of beta-amyloid. (1) can be used in the treatment and prevention of neurodegenerative diseases such as dementia and Alzheimer's disease. The present sequence represents the human APP770 protein which is used in the exemplification of the present invention.

XX Sequence 770 AA;

Query Match 100.0%; Score 4058; DB 23; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRNLNHNHNQNGKWDSPSGTK 60
DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRNLNHNHNQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
QY 241 EADDDDDGDEVEEAPEEATERTTSIATTTTTTSEVEEVREVCSEQATGPC 300
DB 241 EADDDDDGDEVEEAPEEATERTTSIATTTTTTSEVEEVREVCSEQATGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTTEYCVMAVCGSAMSQSLKTTQPLEARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFTTEYCVMAVCGSAMSQSLKTTQPLEARD 360
QY 361 PVKLPPTAASTPDVADKYLETGDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
DB 361 PVKLPPTAASTPDVADKYLETGDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
QY 421 KNLPRADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRLVIYER 540
DB 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFHVRVMDPKAAQIRSQVMTLRLVIYER 540
QY 541 MNQSLSLNYPVAAEEIQDEVELLOKEONYSDOVLANMISEPRISYGNALMPSLTET 600
DB 541 MNQSLSLNYPVAAEEIQDEVELLOKEONYSDOVLANMISEPRISYGNALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEFPVDARPAADRGLTTRPGSLTN 660
DB 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEFPVDARPAADRGLTTRPGSLTN 660
QY 661 IKTEIISEVKMDAEPHRDSEGVHHQKLVFFAEADVGSNKGAIIGLMGVGVVIVITL 720
DB 661 IKTEIISEVKMDAEPHRDSEGVHHQKLVFFAEADVGSNKGAIIGLMGVGVVIVITL 720
QY 721 VMLKKKQVTSIHGGVVEVDAVTPERHLSKMOONGYENPTYKFEQON 770
DB 721 VMLKKKQVTSIHGGVVEVDAVTPERHLSKMOONGYENPTYKFEQON 770

RESULT 13

AAE10650

XX AAE10650 standard; Protein; 772 AA.

AC AAE10650;

XX 10-DEC-2001 (first entry)
XX Human amyloid protein precursor 770-KK (APP770-KK) isoform.
XX Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP770-KK;
XX Alzheimer's disease; Ab; dementia; neurofibrillary tangle; gliosis;
XX amyloid plaque; neuronal loss; proteolytic; neurotropic; neuroprotective.
XX Homo sapiens.
OS Synthetic.
PN GB2357767-A.
XX 04-JUL-2001.
PD 22-SEP-2000; 2000GB-0023315.
PF 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99US-0404133.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX Blenkowski MJ, Gurney M;
PI WPI; 2001-444208/48.
XX N-PSDB; AAD17899.
XX Polypeptide comprising fragments of human aspartyl protease with
XX amyloid precursor protein processing activity and alpha-secretase
XX activity, for identifying modulators useful in treating Alzheimer's
XX disease.
XX Disclosure; Page 148-151; 187pp; English.
XX The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
XX Aspl proteins which lack transmembrane domain or amino terminal
XX domain or cytoplasmic domain and retains alpha-secretase activity
XX and amyloid protein precursor (APP) processing activity. The proteins
XX of the invention are useful for assaying hu-Aspl alpha-secretase
XX activity, which in turn is useful for identifying modulators of
XX hu-Aspl alpha-secretase activity, where modulators that increase
XX hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
XX disease (AD) which causes progressive dementia with consequent
XX formation of amyloid plaques, neurofibrillary tangles, gliosis and
XX neuronal loss. Hu-Aspl protease substrate is useful for assaying
XX hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
XX the substrate under acidic conditions and determining the level of
XX hu-Aspl proteolytic activity. The present sequence is human amyloid
XX protein precursor 770-KK (APP770-KK) isoform. APP770-KK isoform is
XX obtained by the addition of two Lys residues (KK motif) at the
XX C-terminal of APP770 protein.
XX Sequence 772 AA;

Query Match 100.0%; Score 4058; DB 22; Length 772;

Best Local Similarity 100.0%; Pred. No. 5.3e-289;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRNLNHNHNQNGKWDSPSGTK 60
DB 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOQIAMFCGRNLNHNHNQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIONWCKRGKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAEESDNVDSADABEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADABEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPEFYGGCGGNRRNFDEEFCMAVCGSAMQSLSLLKTQOEPLARD 360
Db 301 RAMISRWYFDVTGKCAPEFYGGCGGNRRNFDEEFCMAVCGSAMQSLSLLKTQOEPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
QY 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
Db 541 MNQSLSLYNVPAVEEIQDEVELLQKEQNYSDVLANNISPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLSLDLPWHSFGADSVDPANTENEVEPVDARPAADRLTTPRGSGLTN 660
Db 601 KTTVELLPVNGEFLSLDLPWHSFGADSVDPANTENEVEPVDARPAADRLTTPRGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAITGLMVGVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAITGLMVGVIATVIVITL 720
QY 721 VMLKKKQYTSIHGVEVDEDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGVEVDEDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
RESULT 14
AAE06895
ID .AAE06895 standard; Protein: 772 AA.
AC AAE06895;
XX 23-OCT-2001 (first entry)
XX Human amyloid precursor protein 770-KK (App770-KK) isoform.
DE
XX
KW Human; aspartyl protease; Asp; beta-amyloid precursor protein 770-KK;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nototropic;
KW neuroprotective; antisense therapy; gene therapy; App770-KK; mutant;
KW mutin.
XX
OS Homo sapiens.
OS Synthetic.
OS
PN WO200150829-A2.
XX
PD 19-JUL-2001.
XX
XX 09-MAY-2001; 2001WO-IB00799.
XX
XX 09-MAY-2001; 2001WO-IB00799.
XX
XX (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.

(YANR/) YAN R.
Blenkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
WPI; 2001-483072/52.
N-PSDB; AAD13280.
Novel purified polypeptide comprising fragment of mammalian aspartyl
protease 2, lacking Asp2 transmembrane domain and retaining beta
secretase activity of Asp2 useful for identifying inhibitors of Asp2
activity -
Disclosure: Page 177-180; 185pp; English.
The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
precursor protein (APP) isoforms and their corresponding DNA molecules.
Human aspartyl proteases can act as beta-secretase proteases useful for
treating Alzheimer's disease. APP isoforms are useful for identifying
modulators of amyloid-beta peptide production, for use in designing
therapeutics for the treatment and prevention of Alzheimer's disease,
dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
and neuronal loss. APP isoforms are also used in methods for identifying
inhibitors and modulators of human Asp2 activity. The invention relates
to a method for identifying agents that modulate the activity of human
aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
as a means to screen in cellular assays for the inhibitors of beta- and
gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
polymerase chain reactions (PCR). The probes are useful for detecting
Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
blots. The present sequence is modified human amyloid precursor protein
770-KK (APP770-KK) isoform. APP770-KK isoform is obtained by addition of
two Lys residues (KK motif) at the C-terminal end of APP770 isoform.
XX Sequence 772 AA:
Query Match 100.0%; Score 4058; DB 22; Length 772;
Best Local Similarity 100.0%; Pred. No. 5.3e-289;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLLAATAWATALEVPDGNAGLLAEPIAMFCGRLNMHNVQNGKWDSDPSGK 60
Db 1 MLPGLALLLLAATAWATALEVPDGNAGLLAEPIAMFCGRLNMHNVQNGKWDSDPSGK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITONWCKGRKCKQKTHPHFVPIRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITONWCKGRKCKQKTHPHFVPIRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEETHLHWHTVAKETCEKSTNLHDYGMLLPCCIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCEETHLHWHTVAKETCEKSTNLHDYGMLLPCCIDKFR 180
QY 181 GVEFVCCPLAEESDNVDSADABEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADABEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPEFYGGCGGNRRNFDEEFCMAVCGSAMQSLSLLKTQOEPLARD 360
Db 301 RAMISRWYFDVTGKCAPEFYGGCGGNRRNFDEEFCMAVCGSAMQSLSLLKTQOEPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEEROA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540

Db 481 QAVPRPRHVFNMMLKYYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540
 QY 541 MNQSLSLLYNPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
 Db 541 MNQSLSLLYNPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTET 600
 QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
 QY 661 IKTEEISEVMKDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIGLVMGGVVIATVITL 720
 Db 661 IKTEEISEVMKDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIGLVMGGVVIATVITL 720
 QY 721 VMLKKKYTSTIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFQMQN 770
 Db 721 VMLKKKYTSTIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFQMQN 770

RESULT 15

AAU06624

ID AAU06624 standard; Protein; 772 AA.

XX AAU06624;

XX 24-OCT-2001 (first entry)

XX Human Amyloid precursor protein mutant, APP770-KK.

XX Human: Aspartyl protease; Asp2b; beta-secretase; nootropic;
 KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta; APP770-KK; mutant; muteln.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 358

FT /note= "Encoded by GCC"

FT Misc-difference 771..772
 FT /note= "2 Extra Lys residues added compared to
 wild-type APP770"

XX W0200149098-A2.

XX 12-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00798.

XX 09-MAY-2001; 2001WO-IB00798.

XX (BIEN/) BIENKOWSKI M J.

XX (GURN/) GURNEY M E.

XX (HEIN/) HEINRIKSON R L.

XX (PARO/) PARODI L A.

XX (YANK/) YAN R.

XX Blenkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX WPI; 2001-502549/55.

XX Novel purified polypeptide comprising fragment of mammalian aspartyl
 FT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity

XX Disclosure; Page 177-180; 185pp; English.

XX The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp
 CC proteins and vectors expressing them, and a polypeptide (isoform of

CC amyloid protein precursor (APP) comprising the amino acid sequence of an
 CC APP or its fragment containing an APP cleavage site recognizable by a
 CC mammalian beta-secretase, and further comprising two lysine residues at
 CC the carboxyl terminus of the amino acid sequence of the mammalian APP or
 CC APP fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and
 CC amyloid-beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease.
 CC APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
 CC is associated with increased levels of beta processing is useful in
 CC assays relating to the Alzheimer's research. The expression vector is useful
 CC for recombinantly expressing APP. Nucleic acids that hybridize to
 CC Asp oligonucleotides are useful as probes or primers. The probes are
 CC useful for detecting Hu-Asp nucleic acids in in vitro assays and in
 CC Northern and Southern blots. The present sequence is the human
 CC APP770 mutant, APP770-KK which has 2 extra Lys residues added at
 CC the C-terminus compared to the wild-type APP770. The mutation alters the
 CC specificity of the APP gamma-secretase activity and increases the rate
 CC of processing of the amyloid Abeta peptide.

XX Sequence 772 AA;

Query Match 100.0%; Score 4058; DB 22; Length 772;
 Best Local Similarity 100.0%; Pred. No. 5.3e-289;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVONGKWDSPSGTK 60

Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVONGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCEVYPPELOITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120

Db 61 TCIDTKEGILQYCEVYPPELOITNVVEANQPVTIQNWCKRGRKQCKTHPHFVYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCKFLHOERMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEE 240

Db 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEE 240

QY 241 EADDEDEDDGEVEEEAEPEYEATERTTSTATTTTTTSTESVEEVREVCSQAETGPC 300

Db 241 EADDEDEDDGEVEEEAEPEYEATERTTSTATTTTTTSTESVEEVREVCSQAETGPC 300

QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNEDTEECYCHAVCGSAMSQSLLKTQPEPLARD 360

Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNNEDTEECYCHAVCGSAMSQSLLKTQPEPLARD 360

QY 361 PVKLPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEARQA 420

Db 361 PVKLPTTAASTPDVADKYLETTPGDENEHAHFQKAKERLEAKHRERMSQVMREWEAEARQA 420

QY 421 KNLPRADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480

Db 421 KNLPRADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRRLALENYITAL 480

QY 481 QAVPPRPRHVFNMMLKYYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540

Db 481 QAVPPRPRHVFNMMLKYYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRIYER 540

QY 541 MNQSLSLLYNPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTET 600

Db 541 MNQSLSLLYNPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTET 600

QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660

Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660

QY 661 IKTEEISEVMKDAEFRHDSGYEVHHQKLVFAEDVGSNGKGAIIGLVMGGVVIATVITL 720

|||||
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAVITL 720
QY 721 VMLAKKQYTSIHGCVVEVDAAVTPEERHLSKMOQNGYENPTYKFFEQMON 770
Db 721 VMLAKKQYTSIHGCVVEVDAAVTPEERHLSKMOQNGYENPTYKFFEQMON 770

Search completed: January 28, 2003, 14:13:55
Job time : 45.9063 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 10.0956 seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLAANTARALEV.....KMQQNGYENPTYKFEQMQN 770

Scoring table: BLOSUM62

Gapex 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	9	US-09-785-215-2
2	4058	100.0	770	10	US-09-794-927-55
3	4058	100.0	770	10	US-09-795-847-55
4	4058	100.0	770	10	US-09-794-743-55
5	4058	100.0	770	10	US-09-794-748-55
6	4058	100.0	770	10	US-09-904-987-2
7	4058	100.0	770	10	US-09-794-925-55
8	4058	100.0	770	10	US-09-681-442-55
9	4058	100.0	770	10	US-09-149-718-6
10	4058	100.0	772	10	US-09-794-927-59
11	4058	100.0	772	10	US-09-795-847-59
12	4058	100.0	772	10	US-09-794-743-59
13	4058	100.0	772	10	US-09-794-748-59
14	4058	100.0	772	10	US-09-794-925-59
15	4058	100.0	772	10	US-09-681-442-59
16	3943.5	97.2	751	10	US-09-794-927-57
17	3943.5	97.2	751	10	US-09-795-847-57
18	3943.5	97.2	751	10	US-09-794-743-57
19	3943.5	97.2	751	10	US-09-794-748-57

20	3943.5	97.2	751	10	US-09-794-925-57	Sequence 57, Appl
21	3943.5	97.2	751	10	US-09-681-442-57	Sequence 57, Appl
22	3943.5	97.2	751	10	US-09-149-718-4	Sequence 4, Appl
23	3943.5	97.2	753	10	US-09-794-927-61	Sequence 61, Appl
24	3943.5	97.2	753	10	US-09-795-847-61	Sequence 61, Appl
25	3943.5	97.2	753	10	US-09-794-743-61	Sequence 61, Appl
26	3943.5	97.2	753	10	US-09-794-748-61	Sequence 61, Appl
27	3943.5	97.2	753	10	US-09-794-925-61	Sequence 61, Appl
28	3943.5	97.2	753	10	US-09-681-442-61	Sequence 61, Appl
29	3590.5	88.5	695	10	US-09-794-927-10	Sequence 10, Appl
30	3590.5	88.5	695	10	US-09-795-847-10	Sequence 10, Appl
31	3590.5	88.5	695	10	US-09-794-743-10	Sequence 10, Appl
32	3590.5	88.5	695	10	US-09-794-748-10	Sequence 10, Appl
33	3590.5	88.5	695	10	US-09-794-925-10	Sequence 10, Appl
34	3590.5	88.5	695	10	US-09-681-442-10	Sequence 10, Appl
35	3590.5	88.5	695	10	US-09-149-718-2	Sequence 2, Appl
36	3590.5	88.5	697	10	US-09-794-927-16	Sequence 16, Appl
37	3590.5	88.5	697	10	US-09-795-847-16	Sequence 16, Appl
38	3590.5	88.5	697	10	US-09-794-743-16	Sequence 16, Appl
39	3590.5	88.5	697	10	US-09-794-748-16	Sequence 16, Appl
40	3590.5	88.5	697	10	US-09-794-925-16	Sequence 16, Appl
41	3590.5	88.5	697	10	US-09-681-442-16	Sequence 16, Appl
42	3585.5	88.4	695	10	US-09-794-927-14	Sequence 14, Appl
43	3585.5	88.4	695	10	US-09-795-847-14	Sequence 14, Appl
44	3585.5	88.4	695	10	US-09-794-743-14	Sequence 14, Appl
45	3585.5	88.4	695	10	US-09-794-748-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-785-215-2
; Sequence 2, Application US/09785215
; Publication No. US20020187157A1
; GENERAL INFORMATION:
; APPLICANT: JENSEN, Martin Roland et al.
; TITLE OF INVENTION: NOVEL METHOD FOR DOWN-REGULATION OF AMYLOID
; FILE REFERENCE: 3631-0107P
; CURRENT APPLICATION NUMBER: US/09/785,215
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-215-2

Query Match 100.0%; Score 4058; DB 9; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLHMNVQNGKWDSPSGTK	60
Db	1	MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLHMNVQNGKWDSPSGTK	60
QY	61	TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRCKOCTHPHFVPIYRCLVG	120
Db	61	TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRCKOCTHPHFVPIYRCLVG	120
QY	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSKSTNLHDYGMLLPCGIDKFR	180
Db	121	EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSKSTNLHDYGMLLPCGIDKFR	180
QY	181	GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEVEE	240
Db	181	GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEAEVEEVEE	240
QY	241	EADDEDEDEGDEVEEAEPYEATEERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300
Db	241	EADDEDEDEGDEVEEAEPYEATEERTTTSIATTTTTTTSVEEVVREVCSEQAETGPC	300

;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 55
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-795-847-55

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNQVNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNQVNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIONNCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIONNCKRGRKCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVRCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVRCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
DB 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
DB 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
DB 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600

RESULT 4
US-09-794-743-55
; Sequence 55, Application US/09794743
; Patent No. US20010021391A1

;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Blenkowski, Michael J.
;; APPLICANT: Heinrikson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;; TITLE OF INVENTION: USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/6280BC
;; CURRENT APPLICATION NUMBER: US/09/794,743
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 55
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-743-55

Query Match 100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNQVNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNQVNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIONNCKRGRKCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTIONNCKRGRKCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVRCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVVRCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
DB 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMSQSLLKTTQPEPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
DB 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMEWEAEARQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQDKRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600
DB 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQNYSDVLANMISEPRISYGNALMPSLTET 600

Db 184 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Qy 241 EADDEDEDGDEVEEAEPYEATERTTSIATTTTTTTSVEEVREVCSQAGTGC 300
Db 241 EADDEDEDGDEVEEAEPYEATERTTSIATTTTTTTSVEEVREVCSQAGTGC 300
Qy 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSOSLLKTTQEP 360
Db 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSOSLLKTTQEP 360
Qy 361 PVKLPPTAASTPDVADVKYLETGPDENEHAHFOKAKERLEAKHRMSQVME 420
Db 361 PVKLPPTAASTPDVADVKYLETGPDENEHAHFOKAKERLEAKHRMSQVME 420
Qy 421 KNLPKADKKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRL 480
Db 421 KNLPKADKKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRL 480
Qy 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKAAQIRSQ 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKAAQIRSQ 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQEQNYSDVLANMISEPRISYGNALMPS 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQEQNYSDVLANMISEPRISYGNALMPS 600
Qy 601 KTTVELLPVNGEFSLDLQPHWISFGADSVPAANTEVEPVDARPAADRG 660
Db 601 KTTVELLPVNGEFSLDLQPHWISFGADSVPAANTEVEPVDARPAADRG 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVG 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVG 720
Qy 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQM 770
Db 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQM 770

RESULT 7

US-09-794-925-55
; Sequence 55, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Farodi, Luis A.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280H1
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-55

Query Match

100.0%; Score 4058; DB 10; Length 770;

Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLPGALLLLAAWTAARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVQNGWSDPSG 60
Db 1 MLPGALLLLAAWTAARALEVPTDGNAGLLAEPOIAMFCGRLNMHNVQNGWSDPSG 60
Qy 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFV 120
Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFV 120
Qy 121 EFVSDALLVPDKCKFLHQRMDVVCETHLHWHTVAKETCEKSTNLHDYGM 180
Db 121 EFVSDALLVPDKCKFLHQRMDVVCETHLHWHTVAKETCEKSTNLHDYGM 180
Qy 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEVAEE 240
Db 181 GVEFVCCPLAESDNDVSDADEDDSDVWVGADTDYADGSEDKVVEVAEE 240
Qy 241 EADDEDEDGDEVEEAEPYEATERTTSIATTTTTTTSVEEVREVCSQAGTGC 300
Db 241 EADDEDEDGDEVEEAEPYEATERTTSIATTTTTTTSVEEVREVCSQAGTGC 300
Qy 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSOSLLKTTQEP 360
Db 301 RAMISRWFVTEGKCAPFFYGGCGGNRNFDEEYCMVCGSAMSOSLLKTTQEP 360
Qy 361 PVKLPPTAASTPDVADVKYLETGPDENEHAHFOKAKERLEAKHRMSQVME 420
Db 361 PVKLPPTAASTPDVADVKYLETGPDENEHAHFOKAKERLEAKHRMSQVME 420
Qy 421 KNLPKADKKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRL 480
Db 421 KNLPKADKKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRL 480
Qy 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKAAQIRSQ 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPKAAQIRSQ 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQEQNYSDVLANMISEPRISYGNALMPS 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQEQNYSDVLANMISEPRISYGNALMPS 600
Qy 601 KTTVELLPVNGEFSLDLQPHWISFGADSVPAANTEVEPVDARPAADRG 660
Db 601 KTTVELLPVNGEFSLDLQPHWISFGADSVPAANTEVEPVDARPAADRG 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVG 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVG 720
Qy 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQM 770
Db 721 VMLKKKQYTSIIHGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQM 770

RESULT 8

US-09-681-442-55
; Sequence 55, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13


```

; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-681-442-55

Query Match          100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1  MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGTK 60
Db 1  MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240

QY 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300

QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTTEECVMAVCGSAMSQSLLKTTQEPPLARD 360

QY 361 PVKLPPTAASTPDVADVKYLETFCDENEHAFQKAKERLEAKHRERMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADVKYLETFCDENEHAFQKAKERLEAKHRERMSQVMREWEAEARQA 420

QY 421 KNLPKADKRAVIOHFQEKVESLEQEAANERQQLVETHMARVAMLNDRRLALENYITAL 480
Db 421 KNLPKADKRAVIOHFQEKVESLEQEAANERQQLVETHMARVAMLNDRRLALENYITAL 480

QY 481 QAVPPRPFRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIER 540
Db 481 QAVPPRPFRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIER 540

QY 541 MNGSLSLYNVPAVAEEIQDEVDLELQKQNTSDVLANNISEPRISYNDALMPSLTET 600
Db 541 MNGSLSLYNVPAVAEEIQDEVDLELQKQNTSDVLANNISEPRISYNDALMPSLTET 600

QY 601 KTTVELLPVNGEFLDQFWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDQFWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSGLTN 660

QY 661 IKTEELSEVKMDAEFRHDSGYEVHHOKLVFFAEDVCSNKGALIGLMVGCVIATVITL 720
Db 661 IKTEELSEVKMDAEFRHDSGYEVHHOKLVFFAEDVCSNKGALIGLMVGCVIATVITL 720

QY 721 VMLKKQYTSIHGGVGVVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVGVVDAATPEERHLSKMQQNGYENPTYKFFEQMQN 770
```

```

RESULT 9
US-09-149-718-6
; Sequence 6, Application US/09149718
; Patent No. US20020104104A1
; GENERAL INFORMATION:
; APPLICANT: Dora K. Games, Dale B. Schenk, Lisa C. McConlogue,
; APPLICANT: Peter A. Seubert, and Russell E. Rydel
; TITLE OF INVENTION: Method For Identifying Alzheimer's Disease
; TITLE OF INVENTION: Therapeutics Using Transgenic Animal Models
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/149,718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,487
; FILING DATE:
; APPLICATION NUMBER: 08/480,653
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: ANS101CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-149-718-6

Query Match          100.0%; Score 4058; DB 10; Length 770;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1  MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGTK 60
Db 1  MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRLNMHNQNGKWDSDPSGTK 60

QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVPIYRCLVG 120

QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEAEVEEVEE 240

QY 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVVREVCSEQAETGPC 300

QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTTEECVMAVCGSAMSQSLLKTTQEPPLARD 360
```


Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEPGLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNPVAVEIEQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
Db 541 MNQSLLLYNPVAVEIEQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQONGYENPTYKFEQMQN 770
Db 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQONGYENPTYKFEQMQN 770

RESULT 10

US-09-794-927-59
; Sequence 59, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794, 927
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEQIAIFCGRLLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEQIAIFCGRLLNMHNVQNGKWDSPSGTK 60

QY 61 TCIDTKEGILQYCOBYVPELOITNVVNEANQVPTIQNWCKRGRKCKKTHPHFVPIRCLVG 120
Db 61 TCIDTKEGILQYCOBYVPELOITNVVNEANQVPTIQNWCKRGRKCKKTHPHFVPIRCLVG 120
QY 121 EFVSDALLVPCKFLHQERMDVCEETHLRHWTVAKETCSEKSTNLHDYGMLLPGGIDKFR 180
Db 121 EFVSDALLVPCKFLHQERMDVCEETHLRHWTVAKETCSEKSTNLHDYGMLLPGGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEAEVVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEAEVVEE 240
QY 241 EADDEDDDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDEDDDDGDEVEEAEPEEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEPGLARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMVCGSAMQSLLKTTQEPGLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAHFOKAKERLEAKHRMSQVMREWEAEARQA 420
QY 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
Db 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKAAQIRSQVWTHLRVIYER 540
QY 541 MNQSLLLYNPVAVEIEQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
Db 541 MNQSLLLYNPVAVEIEQDEVDLQKEQNSDDVLANNISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNGKGAIIGLMVGWVIATVIVITL 720
QY 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQONGYENPTYKFEQMQN 770
Db 721 VMLKKQYTSIIHGVVEVDAAVTPPEERHLSKMQONGYENPTYKFEQMQN 770

RESULT 11

US-09-795-847-59
; Sequence 59, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795, 847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416, 901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155, 493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404, 133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOQIAMFCGRNLNMHNQNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOQIAMFCGRNLNMHNQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCEVYPELQITNVVEANQPVTTIQNCKRGRKQCKTHPHFVYRCLVG 120
DB 61 TCIDTREGILQYCEVYPELQITNVVEANQPVTTIQNCKRGRKQCKTHPHFVYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTESVEEVVVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTESVEEVVVCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
DB 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
DB 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
DB 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
QY 541 MNQSLSLYNPVAAVEETQDEVDLQEKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLYNPVAAVEETQDEVDLQEKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660
DB 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFFAEDVGSNKGALIGLMVGWGVVATVIVITL 720
DB 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFFAEDVGSNKGALIGLMVGWGVVATVIVITL 720
QY 721 VMLKKQYTSIIHGVVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMON 770
DB 721 VMLKKQYTSIIHGVVVEVDAATPEERHLSKMQQNGYENPTYKFFEQMON 770
```

RESULT 12

US-09-794-743-59
; Sequence 59, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOQIAMFCGRNLNMHNQNGKWDSPSGTK 60
DB 1 MLPGLALLLAATAARALEVPTDGNAGLLAEPOQIAMFCGRNLNMHNQNGKWDSPSGTK 60
QY 61 TCIDTREGILQYCEVYPELQITNVVEANQPVTTIQNCKRGRKQCKTHPHFVYRCLVG 120
DB 61 TCIDTREGILQYCEVYPELQITNVVEANQPVTTIQNCKRGRKQCKTHPHFVYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCEVTHLHWHITVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWGGADTDYADGSEDKVVEVAEEVEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTESVEEVVVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTESVEEVVVCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
DB 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
DB 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPAKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
DB 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRVDPKKAQIRSQVWTHLRVIYER 540
QY 541 MNQSLSLYNPVAAVEETQDEVDLQEKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
DB 541 MNQSLSLYNPVAAVEETQDEVDLQEKQNYSDVLANMISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660
DB 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTENEVEPVDARPAADRLGTLTRPGSLTN 660
```


Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 13

US-09-794-748-59
; Sequence 59, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHVQNGKWDSPSGTK 60
Qy 61 TCIDTKEGILQYQCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYQCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSKSNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEVAEEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSDKVVVEVAEEVEE 240
Qy 241 EADDEDEDGDEVEEAEPEEATERTTSIAITTTTTTTESEVVEVREYVCSQAETGPC 300
Db 241 EADDEDEDGDEVEEAEPEEATERTTSIAITTTTTTTESEVVEVREYVCSQAETGPC 300
Qy 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSSLLKTTQEPILARD 360
Db 301 RAMISRWFVDTGKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSSLLKTTQEPILARD 360

Qy 361 PVKLPPTTAASTPDVADKYLETGCDENEHAHFOKAKERLEAKHRRMSQVMREWEAEQA 420
Db 361 PVKLPPTTAASTPDVADKYLETGCDENEHAHFOKAKERLEAKHRRMSQVMREWEAEQA 420
Qy 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Qy 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVMVDPKKAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVMVDPKKAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLYNVPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLSLYNVPAVAEEIQDEVDLLOKEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
Qy 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDARPAADRLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKKQYTSIHGCVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMON 770

RESULT 14

US-09-794-925-59
; Sequence 59, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280HI
; CURRENT APPLICATION NUMBER: US/09/794,925
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-59

Query Match 100.0%; Score 4058; DB 10; Length 772;
Best Local Similarity 100.0%; Pred. No. 3.9e-267;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAATAWTAARALEVPTDGNAGLLAEPOIAMFCGRNLNMHVQNGKWDSPSGTK 60
Qy 61 TCIDTKEGILQYQCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYQCEVYPELQITNVVEANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQVPTIQNWCKRGRKCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCFHQRMDVCEVTHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCFHQRMDVCEVTHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Db 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240
Qy 241 EADDDDEDDGDEVEEAEPEEATERTTSTATTTTTTTSVEEVREVSCQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEEATERTTSTATTTTTTTSVEEVREVSCQAETGPC 300
Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLKTTQEPPLARD 360
Qy 361 PVKLPPTAASPDVADVKYLETGPDENEHAHFQAKERLEAKHRMSQVMREWEAEERQA 420
Db 361 PVKLPPTAASPDVADVKYLETGPDENEHAHFQAKERLEAKHRMSQVMREWEAEERQA 420
Qy 421 KNLKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
Db 421 KNLKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480
Qy 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Qy 541 MNQSLSLLYNVPVABEIQDEVELLQKEQNSDDVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLSLLYNVPVABEIQDEVELLQKEQNSDDVLANMISEPRISYGNDAIMPSTET 600
Qy 601 KTTVELLPVNGEFSLDLQPHWSEFGADSVDPANTENEVEPVDARPAADRLTRPGSGLTN 660
Db 601 KTTVELLPVNGEFSLDLQPHWSEFGADSVDPANTENEVEPVDARPAADRLTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGVIATVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

RESULT 15

US-09-681-442-59

; Sequence 59, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FC
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 59

; LENGTH: 772

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-681-442-59

Query Match

100.0%; Score 4058; DB 10; Length 772;

Best Local Similarity 100.0%; Pred. No. 3.9e-267;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEQIANFCGRLLNMHNVQNGKWDSPSGTK 60

Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEQIANFCGRLLNMHNVQNGKWDSPSGTK 60

Qy 61 TCIDTKEGILQYCEVYPELQITNVVEANQVPTIQNWCKRGRKCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQVPTIQNWCKRGRKCKTHPHFVPIYRCLVG 120

Qy 121 EFVSDALLVPDKCFHQRMDVCEVTHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Db 121 EFVSDALLVPDKCFHQRMDVCEVTHLHWHVAKETCSEKSTNLHDYGMLLPCGIDKFR 180

Qy 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Db 181 GVEFVCCPLAEBSDNVSDAEDSDVWVGADTDYADGSEDKVVEVAEEVAEVEE 240

Qy 241 EADDDDEDDGDEVEEAEPEEATERTTSTATTTTTTTSVEEVREVSCQAETGPC 300

Db 241 EADDDDEDDGDEVEEAEPEEATERTTSTATTTTTTTSVEEVREVSCQAETGPC 300

Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLKTTQEPPLARD 360

Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTTEYCMVCGSAMSQSLKTTQEPPLARD 360

Qy 361 PVKLPPTAASPDVADVKYLETGPDENEHAHFQAKERLEAKHRMSQVMREWEAEERQA 420

Db 361 PVKLPPTAASPDVADVKYLETGPDENEHAHFQAKERLEAKHRMSQVMREWEAEERQA 420

Qy 421 KNLKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480

Db 421 KNLKADKAVIQHFOEKVESLEQEAANERQOLVETHMARVEAMLNDRRRLALENYITAL 480

Qy 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Db 481 QAVPPRPRHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Qy 541 MNQSLSLLYNVPVABEIQDEVELLQKEQNSDDVLANMISEPRISYGNDAIMPSTET 600

Db 541 MNQSLSLLYNVPVABEIQDEVELLQKEQNSDDVLANMISEPRISYGNDAIMPSTET 600

Qy 601 KTTVELLPVNGEFSLDLQPHWSEFGADSVDPANTENEVEPVDARPAADRLTRPGSGLTN 660

Db 601 KTTVELLPVNGEFSLDLQPHWSEFGADSVDPANTENEVEPVDARPAADRLTRPGSGLTN 660

Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGVIATVITL 720

Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFAEDVGSNKGAIIGLMVGVIATVITL 720

Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN 770

Search completed: January 28, 2003, 14:40:06

Job time : 13.0956 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 15.9847 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-2

Perfect score: 4058

Sequence: 1 MLPGLALLLLAAWTARALEV.....RMQNGYENPTYKFEQMQN 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4058	100.0	770	1 US-08-133-248-8	Sequence 8, Appli
2	4058	100.0	770	1 US-08-231-940-1	Sequence 1, Appli
3	4058	100.0	770	2 US-08-641-774-1	Sequence 1, Appli
4	4058	100.0	770	2 US-08-104-165-3	Sequence 3, Appli
5	4058	100.0	770	3 US-08-464-250-3	Sequence 3, Appli
6	4058	100.0	770	4 US-08-464-250-3	Sequence 3, Appli
7	4058	100.0	770	4 US-09-548-372D-55	Sequence 55, Appli
8	4058	100.0	770	4 US-09-548-367D-55	Sequence 55, Appli
9	4058	100.0	772	4 US-09-548-372D-59	Sequence 59, Appli
10	4058	100.0	772	4 US-09-548-367D-59	Sequence 59, Appli
11	3943.5	97.2	751	1 US-08-123-702-4	Sequence 4, Appli
12	3943.5	97.2	751	2 US-08-104-165-2	Sequence 2, Appli
13	3943.5	97.2	751	2 US-08-422-333-2	Sequence 2, Appli
14	3943.5	97.2	751	2 US-08-422-333-21	Sequence 21, Appli
15	3943.5	97.2	751	3 US-08-464-250-2	Sequence 2, Appli
16	3943.5	97.2	751	4 US-08-464-250-2	Sequence 2, Appli
17	3943.5	97.2	751	4 US-08-832-867-5	Sequence 5, Appli
18	3943.5	97.2	751	4 US-09-548-372D-57	Sequence 57, Appli
19	3943.5	97.2	751	4 US-09-548-367D-57	Sequence 57, Appli
20	3943.5	97.2	751	6 5187153-2	Patent No. 5187153
21	3943.5	97.2	751	6 5223482-2	Patent No. 5223482
22	3943.5	97.2	753	4 US-09-548-372D-61	Sequence 61, Appli
23	3943.5	97.2	753	4 US-09-548-367D-61	Sequence 61, Appli
24	3937.5	97.0	751	6 5220013-2	Patent No. 5220013
25	3590.5	88.5	695	1 US-08-123-702-2	Sequence 2, Appli
26	3590.5	88.5	695	2 US-08-104-165-1	Sequence 1, Appli
27	3590.5	88.5	695	3 US-08-464-250-1	Sequence 1, Appli

28	3590.5	88.5	695	4 US-08-464-250-1	Sequence 1, Appli
29	3590.5	88.5	695	4 US-09-458-481B-7	Sequence 7, Appli
30	3590.5	88.5	695	4 US-09-458-481B-8	Sequence 8, Appli
31	3590.5	88.5	695	4 US-09-548-372D-10	Sequence 10, Appli
32	3590.5	88.5	695	4 US-09-548-367D-10	Sequence 10, Appli
33	3590.5	88.5	695	6 5218100-2	Patent No. 5218100
34	3590.5	88.5	697	4 US-09-548-372D-16	Sequence 16, Appli
35	3590.5	88.5	697	4 US-09-548-367D-16	Sequence 16, Appli
36	3585.5	88.4	695	4 US-09-548-372D-14	Sequence 14, Appli
37	3585.5	88.4	695	4 US-09-548-367D-14	Sequence 14, Appli
38	3585.5	88.4	697	4 US-09-548-372D-20	Sequence 20, Appli
39	3585.5	88.4	697	4 US-09-548-367D-20	Sequence 20, Appli
40	3584.5	88.3	694	1 US-08-339-152A-18	Sequence 18, Appli
41	3584.5	88.3	694	2 US-08-007-999B-5	Sequence 5, Appli
42	3584.5	88.3	694	2 US-08-689-276A-5	Sequence 5, Appli
43	3582.5	88.3	695	4 US-09-548-372D-12	Sequence 12, Appli
44	3582.5	88.3	695	4 US-09-548-367D-12	Sequence 12, Appli
45	3582.5	88.3	697	4 US-09-548-372D-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-08-133-248-8
; Sequence 8, Application US/08133248
; Patent No. 5525714
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MUTATED FORM OF THE BETA-AMYLOID PRECURSOR
; NUMBER OF SEQUENCES: 8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,248
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-133-248-8

Query Match	100.0%;	Score 4058;	DB 1;	Length 770;
Best Local Similarity	100.0%;	Pred. No. 1.1e-280;		
Matches 770;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEFQIAMFCGRLLNMHNVONGKWDSPSGTK	60	
Db	1	MLPGLALLLLAAWTARALEVPTDGNAGLLAEFQIAMFCGRLLNMHNVONGKWDSPSGTK	60	
Qy	61	TCIDTKEGILQYQCBVPELOITNVVEANQPVITQNWCKRGRKCKTHPHVPIVRCVLG	120	
Db	61	TCIDTKEGILQYQCBVPELOITNVVEANQPVITQNWCKRGRKCKTHPHVPIVRCVLG	120	
Qy	121	EFVSDALLVPDKCKFLHQERMDVCEHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR	180	
Db	121	EFVSDALLVPDKCKFLHQERMDVCEHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR	180	
Qy	181	GVEFFCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE	240	
Db	181	GVEFFCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE	240	
Qy	241	EADDEDDGDEVEEAEEPEEATERTTSTATTTTTTTSVEEVREVSEQAETGCP	300	
Db	241	EADDEDDGDEVEEAEEPEEATERTTSTATTTTTTTSVEEVREVSEQAETGCP	300	
Qy	301	RAMISRWFDTYETGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP	360	
Db	301	RAMISRWFDTYETGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMQSLLKTTQEP	360	


```
Db 301 RAMISRWFDTGKCAPFFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQOEPLARD 360
Qy 361 PVKLPTTAASPTDAVDKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPTTAASPTDAVDKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPKADKKAIVQHFKQKVSLEQEAANERQQLVETHMARVEAMLNDRRLALENVITL 480
Db 421 KNLPKADKKAIVQHFKQKVSLEQEAANERQQLVETHMARVEAMLNDRRLALENVITL 480
Qy 481 QAVPPRRHVFNMKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRRHVFNMKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQYSDVLANMISPRISYGNLALMPSLTET 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQYSDVLANMISPRISYGNLALMPSLTET 600
Qy 601 KTTVELLPVNGEFLSDDLQPMWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLSDDLQPMWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMQN 770

RESULT 2
US-08-231-940-1
; Sequence 1, Application US/08231940
; Patent No. 5550216
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: GELATINASE A INHIBITOR
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,940
; FILING DATE: 25-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-120457
; FILING DATE: 26-APR-1993
; PRIOR APPLICATION NUMBER: JP 6-62129
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: NEIMARK, Sheridan
; REGISTRATION NUMBER: 20,520
; REFERENCE/DOCKET NUMBER: MIYAZAKI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-231-940-1

Query Match 100.0%; Score 4058; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNVONGKWDSDPSGTK 60
Db 1 MLPLGLALLLAATAWTAARALEVPTDGNAGLLAEPQIAMFCGRNLNMHNVONGKWDSDPSGTK 60
Qy 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTTQNWCKRGKCKCKTHPHFVPIPRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNNVEANQPVTTQNWCKRGKCKCKTHPHFVPIPRCLVG 120
Qy 121 EFVSALLVPKCKFLHQERMDVCEVTHLHWHVAVKETSCKSTNLHLDYGMLLPCGIDKFR 180
Db 121 EFVSALLVPKCKFLHQERMDVCEVTHLHWHVAVKETSCKSTNLHLDYGMLLPCGIDKFR 180
Qy 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Db 181 GVEFYCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEE 240
Qy 241 EADDEDEDEDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Db 241 EADDEDEDEDEVEEAEPEYEATERTTSIATTTTTTSTESVEEVREVCSQAETGPC 300
Qy 301 RAMISRWFDTGKCAPFFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQOEPLARD 360
Db 301 RAMISRWFDTGKCAPFFYGGGNNRNFDEEYCMVCGSAMQSLLKTTQOEPLARD 360
Qy 361 PVKLPTTAASPTDAVDKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPTTAASPTDAVDKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPKADKKAIVQHFKQKVSLEQEAANERQQLVETHMARVEAMLNDRRLALENVITL 480
Db 421 KNLPKADKKAIVQHFKQKVSLEQEAANERQQLVETHMARVEAMLNDRRLALENVITL 480
Qy 481 QAVPPRRHVFNMKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Db 481 QAVPPRRHVFNMKKYVRAEQDRQHTLKHFEHVRMVDPKKAAQIRSOVMTHLRVIYER 540
Qy 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQYSDVLANMISPRISYGNLALMPSLTET 600
Db 541 MNQSLSLLYNVPAAVEIQDEVDLQKEQYSDVLANMISPRISYGNLALMPSLTET 600
Qy 601 KTTVELLPVNGEFLSDDLQPMWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLSDDLQPMWHSFGADSVDPANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Qy 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVIVITL 720
Qy 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMQN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMOONGYENPTYKFFEQMQN 770

RESULT 3
US-08-641-774-1
; Sequence 1, Application US/08641774
; Patent No. 5843695
; GENERAL INFORMATION:
; APPLICANT: MIYAZAKI, Kaoru
; TITLE OF INVENTION: BETA-AP DECOMPOSING AGENT
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
```


STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/641,774
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/232,474
FILING DATE: 25-APR-1994
APPLICATION NUMBER: JP 5-122207
FILING DATE: 26-APR-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-51133
FILING DATE: 08-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: NEIMARK, Sheridan
REGISTRATION NUMBER: 20,520
REFERENCE/DOCKET NUMBER: MIYAZAKI-3
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-641-774-1

Query Match 100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.le-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLAANTARALEVPTDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSDPSGK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSDPSGK 60
QY 61 TCIDTREGILQYCEVYPELQINNVANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTREGILQYCEVYPELQINNVANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVFEVCCPLAESDNVDSADAEEDSDVVMGGADTDYADGSEDKVYVEAEVEEVEE 240
DB 181 GVFEVCCPLAESDNVDSADAEEDSDVVMGGADTDYADGSEDKVYVEAEVEEVEE 240
QY 241 EADDEDEDDGVEEAEPEYEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300
DB 241 EADDEDEDDGVEEAEPEYEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300
QY 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQOEPLARD 360
DB 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQOEPLARD 360
QY 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
QY 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
DB 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Query Match 100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.le-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLAANTARALEVPTDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSDPSGK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSDPSGK 60
QY 61 TCIDTREGILQYCEVYPELQINNVANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTREGILQYCEVYPELQINNVANQPVTIQNCKRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVFEVCCPLAESDNVDSADAEEDSDVVMGGADTDYADGSEDKVYVEAEVEEVEE 240
DB 181 GVFEVCCPLAESDNVDSADAEEDSDVVMGGADTDYADGSEDKVYVEAEVEEVEE 240
QY 241 EADDEDEDDGVEEAEPEYEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300
DB 241 EADDEDEDDGVEEAEPEYEATERTSIATTTTTTSTESVEEVVREVCEQATGTC 300
QY 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQOEPLARD 360
DB 301 RAMISRWYFDVTGKCAPFYGGCGNRNFDPEEYCMVCGSAMQSLLKTTQOEPLARD 360
QY 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRMSQVMREWEAEERQA 420
QY 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
DB 421 KNLPKADKAVIQHFQEKVESLQEAANERQQLVETHMARVEAMLNDRRLALENITVAL 480
QY 481 QAVPPRPRHVNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540

Query Match 100.0%; Score 4058; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.le-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLAANTARALEVPTDGNAGLAEPQIAMFCGRNLNHNMVQNGKWDSDPSGK 60


```
Db 1 MLPGLALLLAATAARALEVPTDGNAGLAEPOIAMFCGRNLNHNMQKWDSPSGTK 60
Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Qy 241 EADDDDEDDGDEVEEAEPEYEATERTTSTATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSTATTTTTTSTESVEEVVREVCSEQAETGPC 300
Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD 360
Qy 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Qy 481 QAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRIYER 540
Db 481 QAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRIYER 540
Qy 541 MNQSLSLLYNVPVAAEETQDEVDLQEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLSLLYNVPVAAEETQDEVDLQEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
Qy 601 KTTVELLPVNGEFLDLOPWFHSFGADSPANTENEVEPDVADPAADRLGTLTPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLOPWFHSFGADSPANTENEVEPDVADPAADRLGTLTPGSGLTN 660
Qy 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Db 661 IKTEISEVKMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGWVIATVITL 720
Qy 721 VMLKKQYTSIHGGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770
Db 721 VMLKKQYTSIHGGVVEVDAAVTPEERHLSKMQONGYENPTYKFFEQMQN 770
```

RESULT 5

```
US-08-464-250-3
; Sequence 3, Application US/08464250
; Patent No. 6107542
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTER-HARLIN, Marie-Christine
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-464-250-3
```

Query Match 100.0%; Score 4058; DB 3; Length 770;

Best Local Similarity 100.0%; Pred. No. 1,le-280;

Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MLPGLALLLAATAARALEVPTDGNAGLAEPOIAMFCGRNLNHNMQKWDSPSGTK 60
Db 1 MLPGLALLLAATAARALEVPTDGNAGLAEPOIAMFCGRNLNHNMQKWDSPSGTK 60
Qy 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVITQNWCKRGRKQCKTHPHFVPIYRCLVG 120
Qy 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHWHHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEE 240
Qy 241 EADDDDEDDGDEVEEAEPEYEATERTTSTATTTTTTSTESVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEPEYEATERTTSTATTTTTTSTESVEEVVREVCSEQAETGPC 300
Qy 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECYMAVCGSAMSQSLLKTTQEPPLARD 360
Qy 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADVKYLETPGDENEHAHFQKAKERLEAKHRMSQVMREWEAEARQA 420
Qy 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRKADKAVIQHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
Qy 481 QAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRIYER 540
Db 481 QAVPPRPHVFNMLKKYVRAEQKDRQHTLKHFEHVRVMDPKKAAQIRSQVMTHLRIYER 540
Qy 541 MNQSLSLLYNVPVAAEETQDEVDLQEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLSLLYNVPVAAEETQDEVDLQEQNYSDVLANMISEPRISYGNDAIMPSTLET 600
```


QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDAADRLTTRPGSGLTN 660
QY 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGAIIGLMVGWVVIATVITL 720
Db 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGAIIGLMVGWVVIATVITL 720
QY 721 VMLKKQYTSIHGCVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKQYTSIHGCVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770
RESULT 6
US-08-464-250-3
; Sequence 3, Application US/08464250
; Patent No. 6300540
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; GOATE, Allison Mary
; MULLAN, Michael John
; CHARTIER-HARLIN, Marie-Christine
; OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-464-250-3
Query Match 100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLPGLALLLAAWTAARALEVPTDGNAGLLAEPQIAMFCGRINNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLAAWTAARALEVPTDGNAGLLAEPQIAMFCGRINNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEVPELQITNVVEANQPVTIQNWKGRGRKQCKTHPHFVPIYRCLVG 120

Db 61 TCIDTKEGILQYCOEVPELQITNVVEANQPVTIQNWKGRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCEETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCEETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKPR 180
QY 181 GYEVFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEPE 240
Db 181 GYEVFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVEAEVEPE 240
QY 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTTSVEEVVREVCSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEEAEPEYEATERTTSIATTTTTTTSVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYMAVCSSAMSQSLLKTTQEPPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFDTEECYMAVCSSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDADVKKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDADVKKYLETPGDENEHAHFOKAKERLEAKHRERMSQVMREWEAEARQA 420
QY 421 KNLPRADKXAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPRADKXAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRHVFNNMLKKYVRAEQKDRQHTLKHFEHVMVMDPKKAAQIRSOVMTHLRIYER 540
Db 481 QAVPPRRHVFNNMLKKYVRAEQKDRQHTLKHFEHVMVMDPKKAAQIRSOVMTHLRIYER 540
QY 541 MNQSLLYNVPAAVEIODEVDELLOKEQNYSDOVLANMISEPRISYNDALMPSLTET 600
Db 541 MNQSLLYNVPAAVEIODEVDELLOKEQNYSDOVLANMISEPRISYNDALMPSLTET 600
QY 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDLPQWHSFGADSVDPANTENEVEPVDAADRLTTRPGSGLTN 660
QY 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGAIIGLMVGWVVIATVITL 720
Db 661 IKTEIESEVKMDAEFRHDSGYEVHHQKLVFFAEVDGSKNGAIIGLMVGWVVIATVITL 720
QY 721 VMLKKQYTSIHGCVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770
Db 721 VMLKKQYTSIHGCVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770
RESULT 7
US-09-548-372D-55
; Sequence 55, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-548-372D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKRGKCKTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKRGKCKTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEVAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTTSVEEVVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTTSVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTOEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTOEPLARD 360
QY 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITL 480
DB 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPVPAVEEIQDEVELLQEQNSDDVLANNISPRISYGNDAIMPSTET 600
DB 541 MNQSLSLLYNPVPAVEEIQDEVELLQEQNSDDVLANNISPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEETSEVRMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
DB 661 IKTEETSEVRMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
DB 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 8

US-09-548-372D-55
; Sequence 55, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-55

Query Match 100.0%; Score 4058; DB 4; Length 770;

Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAAMTARALEVPTDGNAGLLAEPOIAMFCGRLLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKRGKCKTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNWKRGKCKTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEVAEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEVEEVAEVEE 240
QY 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTTSVEEVVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPEEATERTTSIATTTTTTTSVEEVVREVCSQAETGPC 300
QY 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTOEPLARD 360
DB 301 RAMISRWYFDVTEGKCAPFFYGGCGGNRNFDTEECMAVCGSAMQSLLKTTOEPLARD 360
QY 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
DB 361 PVKLPTTAASTPDVADKYLETGPDENEHAHFQKAKERLEAKHRERMSQVMREWEAEERQA 420
QY 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITL 480
DB 421 KNLPAKDKKAVIQHFQEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENITL 480
QY 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
DB 481 QAVPPRPRHVFNMKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIER 540
QY 541 MNQSLSLLYNPVPAVEEIQDEVELLQEQNSDDVLANNISPRISYGNDAIMPSTET 600
DB 541 MNQSLSLLYNPVPAVEEIQDEVELLQEQNSDDVLANNISPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLDDLPWHSGADSVPAANTEVEPEVDARPAADRGLTTRPGSGLTN 660
QY 661 IKTEETSEVRMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
DB 661 IKTEETSEVRMDAEFRHDSGYEVHHQKLVFFAEADVGSNKGAIIGLMVGGVVIATVIVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770
DB 721 VMLKKKQYTSIHGGVVEVDAVTPPEERHLSKMQONGYENPTYKFFEQMQN 770

RESULT 9

US-09-548-372D-59
; Sequence 59, Application US/09548372D


```
Patent No. 6420534
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 29915/62801
CURRENT FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: US/09/548,372D
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 60/101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59
LENGTH: 772
TYPE: PRT
ORGANISM: Homo sapiens
US-09-548-372D-59

Query Match      100.0%; Score 4058; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWFYDVTGCKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
DB 301 RAMISRWFYDVTGCKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
DB 361 PVKLPPTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
QY 421 KNLPKADKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRADKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPHVFNMLKYYVRAQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRLVIYER 540
DB 481 QAVPPRPHVFNMLKYYVRAQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRLVIYER 540
QY 541 MNQSLSLYNNPVAEETQDEVELLQEQNYSDVLANMISEPRIYSGNDALMPSLTET 600
DB 541 MNQSLSLYNNPVAEETQDEVELLQEQNYSDVLANMISEPRIYSGNDALMPSLTET 600
QY 601 KTTVELLPVNGEFSLLDLPWHSFGADSVPAANTENEPYDVARPAADRLTTRPGSGLTN 660
DB 601 KTTVELLPVNGEFSLLDLPWHSFGADSVPAANTENEPYDVARPAADRLTTRPGSGLTN 660
QY 661 IKTEISEVKMDAFRHDGSGYEVHVKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
```

```
DB 661 IKTEISEVKMDAFRHDGSGYEVHVKLVFFAEDVGSNKGAIIGLMVGGVIATVITL 720
QY 721 VMLKKKQYTSIIHGVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770
DB 721 VMLKKKQYTSIIHGVVEVDAVTPERHLSKMQONGYENPTYKFFEQMON 770

RESULT 10
US-09-548-367D-59
Sequence 59, Application US/09548367D
Patent No. 640698
GENERAL INFORMATION:
APPLICANT: GURNEY ET AL.
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 29915/6280H
CURRENT FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: US 60/155,493
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 09/404,133
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: PCT/US99/20881
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: US 60/101,594
PRIOR FILING DATE: 1998-09-24
NUMBER OF SEQ ID NOS: 73
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59
LENGTH: 772
TYPE: PRT
ORGANISM: Homo sapiens
US-09-548-367D-59

Query Match      100.0%; Score 4058; DB 4; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.1e-280;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
DB 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVIPYRCLVG 120
DB 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTIQNCKRGRKCKOCTHPHFVIPYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE 240
DB 181 GVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVAEEVEE 240
QY 241 EADDDDEDDGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEPYEATERTTSIATTTTTTSTESVEEVVREVCSQAETGPC 300
QY 301 RAMISRWFYDVTGCKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
DB 301 RAMISRWFYDVTGCKCAPFFYGGCGGNRNFDTEECYCMVCGSAMSQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
DB 361 PVKLPPTAASTPDADVKYLETGPDENEHAHFQKAKERLEAKHRMSQVWREWEAERQA 420
QY 421 KNLPKADKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPKADKAVIOHFQEKVESLEQEAANERQOLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPHVFNMLKYYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRSQVMTLRLVIYER 540
```


Db 481 QAVPPRRHVNMLKKYVRAEQDKRQHTLKHEHVRMVDPKAAQIRSOVMTHLVIYER 540
QY 541 MNQSLLLYNVPAVAEIQDEYDELLQEQNYSDVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLLLYNVPAVAEIQDEYDELLQEQNYSDVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDQOPHSHFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDQOPHSHFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADYGSNKGALIGLMVGVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADYGSNKGALIGLMVGVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMGN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMGN 770

RESULT 11
US-08-123-702-4
; Sequence 4, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Vermuri, B.
; APPLICANT: Wei, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: 1201 West Peachtree Street
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 17-SEPT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-123-702-4

Query Match 97.2%; Score 3943.5; DB 1; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

QY 1: MLPGLALLLAANTARALEVPTDGNAGLLAEPTAMFCGRNLNMHVMYQNGKWDSPSGTK 60
Db 1 MLPGLALLLAANTARALEVPTDGNAGLLAEPTAMFCGRNLNMHVMYQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVITQNCWKGRKQCKTHPHFVYRCLVG 120

Db 61 TCIDTKEGILQYCEVYPELQITNVVEANQPVITQNCWKGRKQCKTHPHFVYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSNLHDYGMLLPCGDKPFR 180
Db 121 EFVSDALLVPDKCKFLHQERMDVCETHLHHHTVAKETCSEKSNLHDYGMLLPCGDKPFR 180
QY 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240
Db 181 GVEFVCCPLAEESDNVDSADAEEDSDVWVGADTDYADGSDKVVEVAEEVAEVEE 240
QY 241 EADDEDEDGDEVEEAEPYEEATERITTSIATTTTTTTESEVEEVVREYVCSQAETGPC 300
Db 241 EADDEDEDGDEVEEAEPYEEATERITTSIATTTTTTTESEVEEVVREYVCSQAETGPC 300
QY 301 RAMLSRWYFDVTGCKCAPFYGGCGNNRNFDEEYCMVCGSAMQSLLKTTQEPPLARD 360
Db 301 RAMLSRWYFDVTGCKCAPFYGGCGNNRNFDEEYCMVCGSAMQSLLKTTQEPPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETGPDENEHAFQAKERLEAKHRRMSQVMREWEAEROA 420
Db 361 PVKLPPTAASTPDVADKYLETGPDENEHAFQAKERLEAKHRRMSQVMREWEAEROA 420
QY 421 KNLPAKADKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPAKADKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRRHVNMLKKYVRAEQDKRQHTLKHEHVRMVDPKAAQIRSOVMTHLVIYER 540
Db 481 QAVPPRRHVNMLKKYVRAEQDKRQHTLKHEHVRMVDPKAAQIRSOVMTHLVIYER 540
QY 541 MNQSLLLYNVPAVAEIQDEYDELLQEQNYSDVLANMISEPRISYGNDAIMPSTET 600
Db 541 MNQSLLLYNVPAVAEIQDEYDELLQEQNYSDVLANMISEPRISYGNDAIMPSTET 600
QY 601 KTTVELLPVNGEFLDQOPHSHFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
Db 601 KTTVELLPVNGEFLDQOPHSHFGADSVPAANTENEVEPVDARPAADRLTTRPGSGLTN 660
QY 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADYGSNKGALIGLMVGVIATVITL 720
Db 661 IKTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEADYGSNKGALIGLMVGVIATVITL 720
QY 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMGN 770
Db 721 VMLKKKQYTSIHGGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQMGN 770

RESULT 12
US-08-104-165-2
; Sequence 2, Application US/08104165
; Patent No. 5877015
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,165

;
; FILING DATE: 21-JAN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-104-165-2

Query Match 97.2%; Score 3943.5; DB 2; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMSSQLLKTTOEPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMSSQLLKTTOEPLARD 360
QY 361 PVKLPPTAASTPDVADKYLETPGDENEHAFQKAKERLEAKHRRMSQVMREWEAEARQA 420
Db 361 PVKLPPTAASTPDVADKYLETPGDENEHAFQKAKERLEAKHRRMSQVMREWEAEARQA 420
QY 421 KNLPKADKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
Db 421 KNLPKADKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNLKYYVRAEKQDRQHTLKHFEHVMWDPKKAQIRSQVMTHLRVIYER 540
Db 481 QAVPPRPRHVFNLKYYVRAEKQDRQHTLKHFEHVMWDPKKAQIRSQVMTHLRVIYER 540
QY 541 MNQSLLYNVPAVAEEIODEVELLQKEQNSDDVLANNKISEPRISYGNDAIMPSTLET 600
Db 541 MNQSLLYNVPAVAEEIODEVELLQKEQNSDDVLANNKISEPRISYGNDAIMPSTLET 600
QY 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTEVEPEVDARPAADRLGLTTRPGSLTN 660
Db 601 KTTVELLPVNGEFLDLDLPWHSFGADSVPAANTEVEPEVDARPAADRLGLTTRPGSLTN 660
QY 661 IKTEELSEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGWGIATVIVITL 720
Db 661 IKTEELSEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGWGIATVIVITL 720
```

```
Db 642 IKTEELSEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGWGIATVIVITL 701
QY 721 VMLKKKQYTSIIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 770
Db 702 VMLKKKQYTSIIHHGVVEVDAAVTPPEERHLSKMQQNGYENPTYKFFEQMON 751

RESULT 13
US-08-422-333-2
; Sequence 2, Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
; APPLICANT: CORDELL, Barbara L.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
; TITLE OF INVENTION: THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scios, Inc.
; STREET: 2450 Bayshore Parkway
; CITY: Mountain View
; STATE: CA
; COUNTRY: USA
; ZIP: 94043
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-422-333-2
```

Query Match 97.2%; Score 3943.5; DB 2; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

```
QY 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
Db 1 MLPGLALLLLAAWTARALEVPTDGNAGLLAEPOIAMFCGRNLNMHNVQNGKWDSPSGTK 60
QY 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKQCKTHPHFVPIYRCLVG 120
Db 61 TCIDTKEGILQYCOEYVPELQITNVVEANQPVTTQNMCKRGKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHVTAKETSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
Db 181 GVEFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSEDKVVEVAEEVEAEVEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
Db 241 EADDDDEDDGDEVEEAEEPEEATERTTSIATTTTTTSTESVEEVVRCVSEQAETGPC 300
QY 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMSSQLLKTTOEPLARD 360
Db 301 RAMISRWFYDVTGKCAPFFYGGCGGNRNFTTEYCMVCGSAMSSQLLKTTOEPLARD 360
```



```
Db 301 RAMISRWYDFVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSA----- 344
Qy 361 PVKLPPTAASTPDADVKYLETGPDENEHAFQAKERLEAKHRERMSQVMREWEAEARQA 420
Db 345 ---IPTTAASTPDADVKYLETGPDENEHAFQAKERLEAKHRERMSQVMREWEAEARQA 401
Qy 421 KNLPKADKAVIQHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 402 KNLPKADKAVIQHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 461
Qy 481 QAVPPRRHVFNMUKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Db 462 QAVPPRRHVFNMUKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 521
Qy 541 MNQSLSLLYNPAVAEETQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTET 600
Db 522 MNQSLSLLYNPAVAEETQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTET 581
Qy 601 KTTVELLPVNGEFSLLDLPWHSEFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 582 KTTVELLPVNGEFSLLDLPWHSEFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 641
Qy 661 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 720
Db 642 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 701
Qy 721 VMLKKKQYTSIIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 702 VMLKKKQYTSIIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 751

RESULT 14
US-08-422-333-21
; Sequence 21. Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
; APPLICANT: CORDELL, Barbara L.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
; THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIOS, Inc.
; STREET: 2450 Bayshore Parkway
; CITY: Mountain View
; STATE: CA
; COUNTRY: USA
; ZIP: 94043
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-422-333-21

Query Match 97.28; Score 3943.5; DB 2: Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
```

```
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;
Qy 1 MLPGIALLLLAAWTARALEVPTDGNAGLLABEQIAMFCGRLNMHNMVONGKWDSPSGTK 60
Db 1 MLPGIALLLLAAWTARALEVPTDGNAGLLABEQIAMFCGRLNMHNMVONGKWDSPSGTK 60
Qy 61 TCIDTKEGILQYCOEYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
Db 61 TCIDTKEGILQYCOEYPELQITNVVEANQPVTIQNWCKRGRKQCKTHPHFVIPYRCLVG 120
Qy 121 EFVSDALLVPDKCKFLHQBQMDVCECHLHHWTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Db 121 EFVSDALLVPDKCKFLHQBQMDVCECHLHHWTVAKETCEKSTNLHDYGMLLPCGIDKFR 180
Qy 181 GVEFFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSDKVVEVAEEVEAEVEEE 240
Db 181 GVEFFVCCPLAESDNVDSADAEEDSDVMWGGADTDYADGSDKVVEVAEEVEAEVEEE 240
Qy 241 EADDEDEDEGDEVEEAEPEYEAEATERSTSTATTTTTTSTSTSTSTSTSTSTSTSTST 300
Db 241 EADDEDEDEGDEVEEAEPEYEAEATERSTSTATTTTTTSTSTSTSTSTSTSTSTSTST 300
Qy 301 RAMISRWYDFVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMSQSLLKTTQEPILAR 360
Db 301 RAMISRWYDFVTEGKCAPFFYGGCGGNRNFTDEEYCMVCGSAMSQSLLKTTQEPILAR 360
Qy 361 PVKLPPTAASTPDADVKYLETGPDENEHAFQAKERLEAKHRERMSQVMREWEAEARQA 420
Db 345 ---IPTTAASTPDADVKYLETGPDENEHAFQAKERLEAKHRERMSQVMREWEAEARQA 401
Qy 421 KNLPKADKAVIQHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 480
Db 402 KNLPKADKAVIQHFQKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITL 461
Qy 481 QAVPPRRHVFNMUKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 540
Db 462 QAVPPRRHVFNMUKKYVRAEQDKRQHTLKHFEHVRMVDPKKAAQIRSQVMTHLRVIYER 521
Qy 541 MNQSLSLLYNPAVAEETQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTET 600
Db 522 MNQSLSLLYNPAVAEETQDEVDLQKQNYSDVLANMISEPRISYGNDAIMPSTET 581
Qy 601 KTTVELLPVNGEFSLLDLPWHSEFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 660
Db 582 KTTVELLPVNGEFSLLDLPWHSEFGADSVPAANTENEVEPVDARPAADRGLTTRPGSGLTN 641
Qy 661 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 720
Db 642 IKTEISEVMKDAEFRHDSGYEVHHOKLVFFAEDVGSNKGAIIGLMVGGVVIATVITL 701
Qy 721 VMLKKKQYTSIIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 770
Db 702 VMLKKKQYTSIIHGGVVEVDAVTPPEERHLSKMQQNGYENPTYKFFEQMQN 751

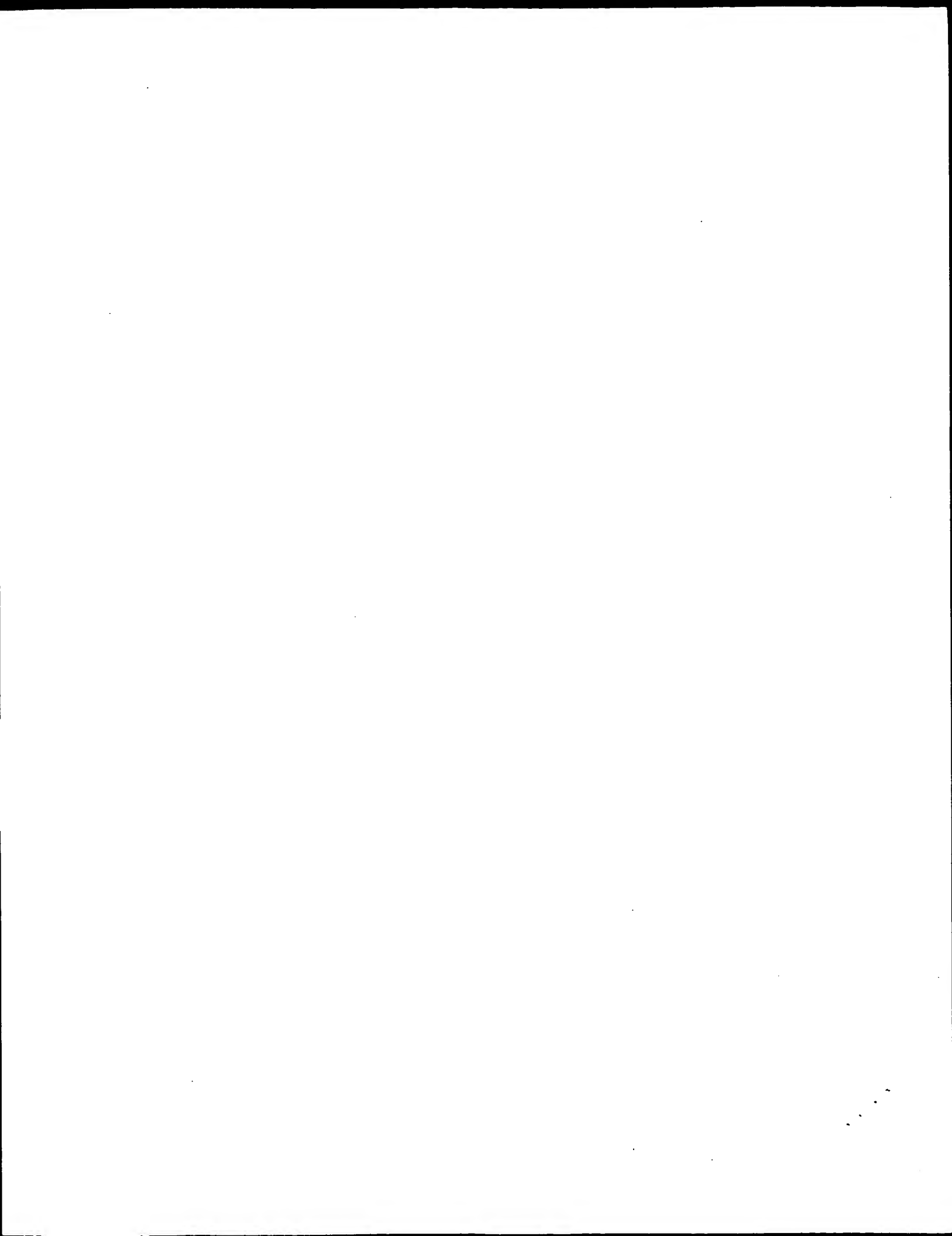
RESULT 15
US-08-464-250-2
; Sequence 2. Application US/08464250
; Patent No. 6107542
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; APPLICANT: OWEN, Michael John
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
```


ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,250
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/104,165
FILING DATE: 21-JAN-1992
APPLICATION NUMBER: 9101307.8
FILING DATE: 21-JAN-1991
APPLICATION NUMBER: 9118445.7
FILING DATE: 28-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 16163-000100
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-464-250-2

Search completed: January 28, 2003, 14:21:20
Job time : 18.9847 secs

Query Match 97.2%; Score 3943.5; DB 3; Length 751;
Best Local Similarity 97.4%; Pred. No. 1.4e-272;
Matches 750; Conservative 1; Mismatches 0; Indels 19; Gaps 1;

QY 1 MLPLGALLLAATARALETPTDGNAGLLAEQIAMFCRLNMHNMVQNGKWDSDPSGTK 60
DB 1 MLPLGALLLAATARALETPTDGNAGLLAEQIAMFCRLNMHNMVQNGKWDSDPSGTK 60
QY 61 TCIDTKEGILQYCOEYPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120
DB 61 TCIDTKEGILQYCOEYPELQITNVVEANQPVTIONWCKRGRKQCKTHPHFVPIYRCLVG 120
QY 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
DB 121 EFVSDALLVPDKCKFLHQRMDVCETHLHWHTVAKETCSEKSTNLHDYGMLLPCGIDKFR 180
QY 181 GFVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEEE 240
DB 181 GFVEFVCCPLAESDNVDSADAEEDSDVWVGADTDYADGSEDKVVEVAEEVEAEVEEE 240
QY 241 EADDDDEDDGDEVEEAEEPEEATERTSTATTTTTTSEVEEVVREVCSEQAETGPC 300
DB 241 EADDDDEDDGDEVEEAEEPEEATERTSTATTTTTTSEVEEVVREVCSEQAETGPC 300
QY 301 RAMISRWYFDVTGKCAPFFYCGCGNRNNDTEECVMAVCGSAMQSLLKTTQEP LARD 360
DB 301 RAMISRWYFDVTGKCAPFFYCGCGNRNNDTEECVMAVCGSAMQSLLKTTQEP LARD 360
QY 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
DB 361 PVKLPTTAASTPDAVDKYLETPGDENEHAHFQAKERLEAKHRERMSQVMREWEAEARQA 420
QY 421 KNLPRADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
DB 421 KNLPRADKKAVIQHFOEKVESLEQEAANERQQLVETHMARVEAMLNDRRLALENYITAL 480
QY 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPPKAAQIRSQVTHLRVIYER 540
DB 481 QAVPPRPRHVFNMLKKYVRAEQDKRQHTLKHFEHVRMVDPPKAAQIRSQVTHLRVIYER 540
QY 541 MNQSLSLLYNPAVAEEIQDEYDELLQEQNYSDVLANMISEPRISYGNALMPSLTET 600



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 6.63425 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCWMLVLFVATMSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1435	100.0	253	1 UJHU	major prion protei
2	1427	99.4	253	2 I61847	major prion protei
3	1427	99.4	253	2 S53635	prion protein - si
4	1427	99.4	253	2 S53614	major prion protei
5	1422	99.1	253	2 S53617	major prion protei
6	1421	99.0	253	2 I37032	major prion protei
7	1409	98.2	253	2 S53616	major prion protei
8	1404	97.8	253	2 S53618	major prion protei
9	1401	97.6	253	2 I84423	major prion protei
10	1401	97.6	253	2 S71055	major prion protei
11	1400	97.6	253	2 S53619	major prion protei
12	1396	97.3	253	2 S53620	major prion protei
13	1396	97.3	253	2 S53623	major prion protei
14	1396	97.3	253	2 S53624	major prion protei
15	1396	97.3	253	2 S53625	major prion protei
16	1386.5	96.6	260	2 S53629	major prion protei
17	1380.5	96.2	252	2 S53634	major prion protei
18	1374.5	95.8	252	2 I61848	major prion protei
19	1370.5	95.5	252	2 S53631	major prion protei
20	1332	92.8	241	2 S71048	major prion protei
21	1331.5	92.8	264	2 A54330	major prion protei
22	1329	92.6	241	2 S71056	major prion protei
23	1329	92.6	245	2 S71045	major prion protei
24	1327	92.5	245	2 S53627	major prion protei
25	1323.5	92.2	256	2 A34759	prion protein - Ch
26	1323.5	92.2	256	2 JU0268	major prion protei
27	1319.5	92.0	254	2 B34759	major prion protei
28	1316.5	91.7	256	2 A54281	major prion protei
29	1311.5	91.4	256	2 S37149	prion protein - go

30	1311.5	91.4	264	2 S37137	prion protein - gr
31	1308.5	91.2	254	1 UJHYIH	major prion prp-Sc
32	1300.5	90.6	239	2 S53633	major prion protei
33	1299	90.5	257	2 J01900	major prion protei
34	1297	90.4	257	2 A23544	major prion prp27-
35	1296.5	90.3	254	2 A23544	major prion protei
36	1280.5	89.2	252	2 JG6175	prion protein - ra
37	1266.5	88.3	232	2 S71041	major prion protei
38	1183.5	82.5	226	2 A53892	prion-related prot
39	430.5	30.0	273	2 A46280	prion protein - ch
40	422.5	29.4	267	2 A37372	prion protein homo
41	421.5	29.4	267	1 UJCH	major prion protei
42	189	13.2	346	1 S35500	heterogeneous ribo
43	173	12.1	465	1 S01820	glycine-rich cell
44	172.5	12.0	200	2 S01334	glycine-rich prote
45	172	12.0	526	1 KRBOVI	keratin, 54K type

ALIGNMENTS

RESULT 1

UJHU

major prion protein precursor - human

N:Alternate names: 11K amyloid protein; 27-30K sialoglycoprotein; PrP 27-30; PrP 33-3

C:Species: Homo sapiens (man)

C>Date: 25-Oct-1987 #sequence_revision 12-Apr-1996 #text_change 16-Jun-2000

C:Accession: A24173; A40372; A05017; S14078; I54322; I68597; I58135; I59184; I796333;

R:Kretschmar, H.A.; Stowring, L.E.; Westaway, D.; Stubblebine, W.H.; Prusiner, S.B.; DNA 5, 315-324, 1986

A:Title: Molecular cloning of a human prion protein cDNA.

A:Reference number: A24173; MUID:86300093; PMID:3755672

A:Accession: A24173

A:Molecule type: mRNA

A:Residues: 1-253 <KRE>

A:Cross-references: GB:M13899; NID:g190467; PIDN:AAA60182.1; PID:g190468

R:Puckett, C.; Concannon, P.; Casey, C.; Hood, L.

Am. J. Hum. Genet. 49, 320-329, 1991

A:Title: Genomic structure of the human prion protein gene.

A:Reference number: A40372; MUID:91328137; PMID:1678248

A:Accession: A40372

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-80,89-253 <PUC>

A:Cross-references: GB:X83416; NID:g747846; PIDN:CAA58442.1; PID:g747847

A>Note: the deletion may be a polymorphism; the alternative deletion of 82-89 could n

R:Liao, Y.C.J.; Lebo, R.V.; Clawson, G.A.; Smuckler, E.A.

Science 233, 364-367, 1986

A:Reference number: A05017; MUID:86261778; PMID:3014653

A:Accession: A05017

A:Molecule type: mRNA

A:Residues: 8-117,119-253 <LTA>

A:Cross-references: GB:D00015; NID:g220015; PIDN:BAA00011.1; PID:g220016; GB:M13667;

R:Tagliavini, F.; Prelli, F.; Ghiso, J.; Bugiani, O.; Serban, D.; Prusiner, S.B.; Far

EMBO J. 10, 513-519, 1991

A:Title: Amyloid protein of Gerstmann-Sträussler-Scheinker disease (Indiana kindred)

A:Reference number: S14078; MUID:91160504; PMID:1672107

A:Accession: S14078

A:Molecule type: protein

A:Residues: 58-72, 'X', 74-76, 'XX', 79, 'XXX', 83-86:111-128, 'V', 130-150 <TAG>

R:Diedrich, J.F.; Knopman, D.S.; List, J.F.; Olson, K.; Frey, W.H.

Hum. Mol. Genet. 1, 443-444, 1992

A:Title: Deletion in the prion protein gene in a demented patient.

A:Reference number: I54322; MUID:93250789; PMID:1363802

A:Accession: I54322

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 9-83,92-240 <RES>

A:Cross-references: GB:M81929; NID:g190517; PIDN:AAB59442.1; PID:g190518

A:Accession: I68597

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 8-240 <RE3>


```

A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-253 <RES>
A:Cross-references: EMBL:U15039; NID:g609303; PIDN:AAA68632.1; PID:g609304
R:Schatztl, H.M.,
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71060
A:Molecule type: DNA
A:Residues: 1-253 <SCW>
A:Cross-references: EMBL:U08296; NID:g474350; PIDN:AAC50085.1; PID:g474351
R:Schatztl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53615
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-210,'R',212-253 <SCH>
A:Cross-references: EMBL:U08296
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; tra
Query Match 99.4%; Score 1427; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 4.8e-109;
Matches 251; Conservative 2; Mismatches 0; Indels 0; Gaps
QY 1 MANLGCWMLVLFVATWSDLGLCKRKPKPGGWNTGGSRYPGGSPGNGNRYPPQGSGGWGP 60
Db 1 MANLGCWMLVLFVATWSDLGLCKRKPKPGGWNTGGSRYPGGSPGNGNRYPPQGSGGWGP 60
QY 61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGTTHSQNNKPSKPTNKKHMAGAAAAG 120
Db 61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGTTHSQNNKPSKPTNKKHMAGAAAAG 120
QY 121 VVGLGGLYMLGSAMSRPIIHGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
Db 121 VVGLGGLYMLGSAMSRPIIHGSDYEDRYRENHRYPNQVYRPMDOYSSQNNFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDKMERVVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDKMERVVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253
RESULT 3
S53635
prion protein - siamang
C:Species: Hylobates syndactylus (siamang)
C:Date: 15-Jul-1995 #sequence_revision 19-Apr-1996 #text_change 13-Aug-1999
C:Accession: S53635
R:Schatztl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53635
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08308; NID:g474374; PIDN:AAC50096.1; PID:g474375
A:Note: the source was designated as Synphalangus syndactylus
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, Apr
C:Superfamily: major prion protein
Query Match 99.4%; Score 1427; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 4.8e-109;
Matches 251; Conservative 2; Mismatches 0; Indels 0; Gaps
QY 1 MANLGCWMLVLFVATWSDLGLCKRKPKPGGWNTGGSRYPGGSPGNGNRYPPQGSGGWGP 60

```



```

Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 4
S53614
major prion protein - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53614; S71049
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53614
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08300
R:Schaezel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71050
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08300; NID:g474358; PIDN:AAC50089.1; PID:g474359
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane
protein

Query Match 99.4%; Score 1427; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 4.8e-109;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 5
S53617
major prion protein - common gibbon
C:Species: Hylobates lar (common gibbon, white-handed gibbon)

```

```

C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53617; S71050
R:Schaezel, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53617
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08299
R:Schaezel, H.M.
submitted to the EMBL Data Library, April 1994
A:Reference number: S71041
A:Accession: S71050
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08299; NID:g474356; PIDN:AAC50088.1; PID:g474357
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane
protein

Query Match 99.1%; Score 1422; DB 2; Length 253;
Best Local Similarity 98.8%; Pred. No. 1.2e-108;
Matches 250; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNRFVHDCV 180
QY 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTKGENFTETDVKMVRVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 6
I37032
major prion protein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999
C:Accession: I37032
R:Cervenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
A:Title: Infectious amyloid precursor gene sequences in primates used for experimenta
A:Reference number: I36907; MUID:95083661; PMID:7991600
A:Accession: I37032
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-253 <RES>
A:Cross-references: EMBL:U15166; NID:g563208; PIDN:AAA68633.1; PID:g563209
C:Superfamily: major prion protein

Query Match 99.0%; Score 1421; DB 2; Length 253;
Best Local Similarity 99.2%; Pred. No. 1.5e-108;
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
Db 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNKHMAGAAAAGA 120

```


C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.6%; Score 1401; DB 2; Length 253;
Best Local Similarity 96.4%; Pred. No. 6.3e-107;
Matches 244; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 10

S71055

major prion protein - pig-tailed macaque

C:Species: Macaca nemestrina (pig-tailed macaque)

C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999

C:Accession: S71055; S53626

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71055

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08306; NID:9474370; PIDN:AAC50094.1; PID:9474371

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53614; MUID:95139066; PMID:7837269

A:Accession: S53626

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 8-210, R', 212-247 <SCW>

A:Cross-references: EMBL:U08306

C:Superfamily: major prion protein

C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.6%; Score 1401; DB 2; Length 253;
Best Local Similarity 96.4%; Pred. No. 6.3e-107;
Matches 244; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240

QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 11

S53619

major prion protein - Presbytis francoisi

C:Species: Presbytis francoisi

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C:Accession: S53619; S71057

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53614; MUID:95139066; PMID:7837269

A:Accession: S53619

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08302

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71057

A:Molecule type: DNA

A:Residues: 1-210, 'E', 212-253 <SCW>

A:Cross-references: EMBL:U08302; NID:gl396067; PIDN:AAB03105.1; PID:gl396068

C:Superfamily: major prion protein

C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match

97.6%; Score 1400; DB 2; Length 253;

Best Local Similarity 96.4%; Pred. No. 7.6e-107;

Matches 244; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGQGP 60
QY 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
Db 61 HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP HGGGQGP 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240

QY 241 ILLISFLIFLIVG 253

Db 241 ILLISFLIFLIVG 253

RESULT 12

S53620

major prion protein - hamadryas baboon

C:Species: Papio hamadryas (hamadryas baboon)

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999

C:Accession: S53620; S71058

R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.

J. Mol. Biol. 245, 362-374, 1995

A:Title: Prion protein gene variation among primates.

A:Reference number: S53614; MUID:95139066; PMID:7837269

A:Accession: S53620

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-253 <SCH>

A:Cross-references: EMBL:U08294

R:Schatzl, H.M.

submitted to the EMBL Data Library, April 1994

A:Reference number: S71041

A:Accession: S71058
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08294; NID:g474346; PIDN:AAC50083.1; PID:g474347
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.3%; Score 1396; DB 2; Length 253;
Best Local Similarity 96.0%; Pred. No. 1.6e-106;
Matches 243; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||
Db 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||

QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
Db 121 VVGGLGYMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||

QY 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253
|||||

RESULT 13
S53623
major prion protein - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53623; S71052
R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53623
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08298
R:Schatzl, H.M.
A:Reference number: S71041
submitted to the EMBL Data Library, April 1994
A:Accession: S71052
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08298; NID:g474354; PIDN:AAC50087.1; PID:g474355
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.3%; Score 1396; DB 2; Length 253;
Best Local Similarity 96.0%; Pred. No. 1.6e-106;
Matches 243; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||
Db 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||

QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
Db 121 VVGGLGYMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||

QY 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253
|||||

RESULT 14
S53624
major prion protein - stump-tailed macaque
C:Species: Macaca arctoides (stump-tailed macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53624; S71051
R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53624
A:Molecule type: DNA
A:Residues: 1-253 <SCH>
A:Cross-references: EMBL:U08311
R:Schatzl, H.M.
A:Reference number: S71041
submitted to the EMBL Data Library, April 1994
A:Accession: S71051
A:Molecule type: DNA
A:Residues: 1-210,'E',212-253 <SCW>
A:Cross-references: EMBL:U08311; NID:g475583; PIDN:AAC50099.1; PID:g475584
C:Superfamily: major prion protein
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 97.3%; Score 1396; DB 2; Length 253;
Best Local Similarity 96.0%; Pred. No. 1.6e-106;
Matches 243; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||
Db 1 MANLGCMMLVLFVATWSDLGLCKRKPGGWNTGSGRYPGQSPGNGRYPPOGGGGWGQP 60
|||||

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKPSKPTNMKHMAGAAAAGA 120
|||||

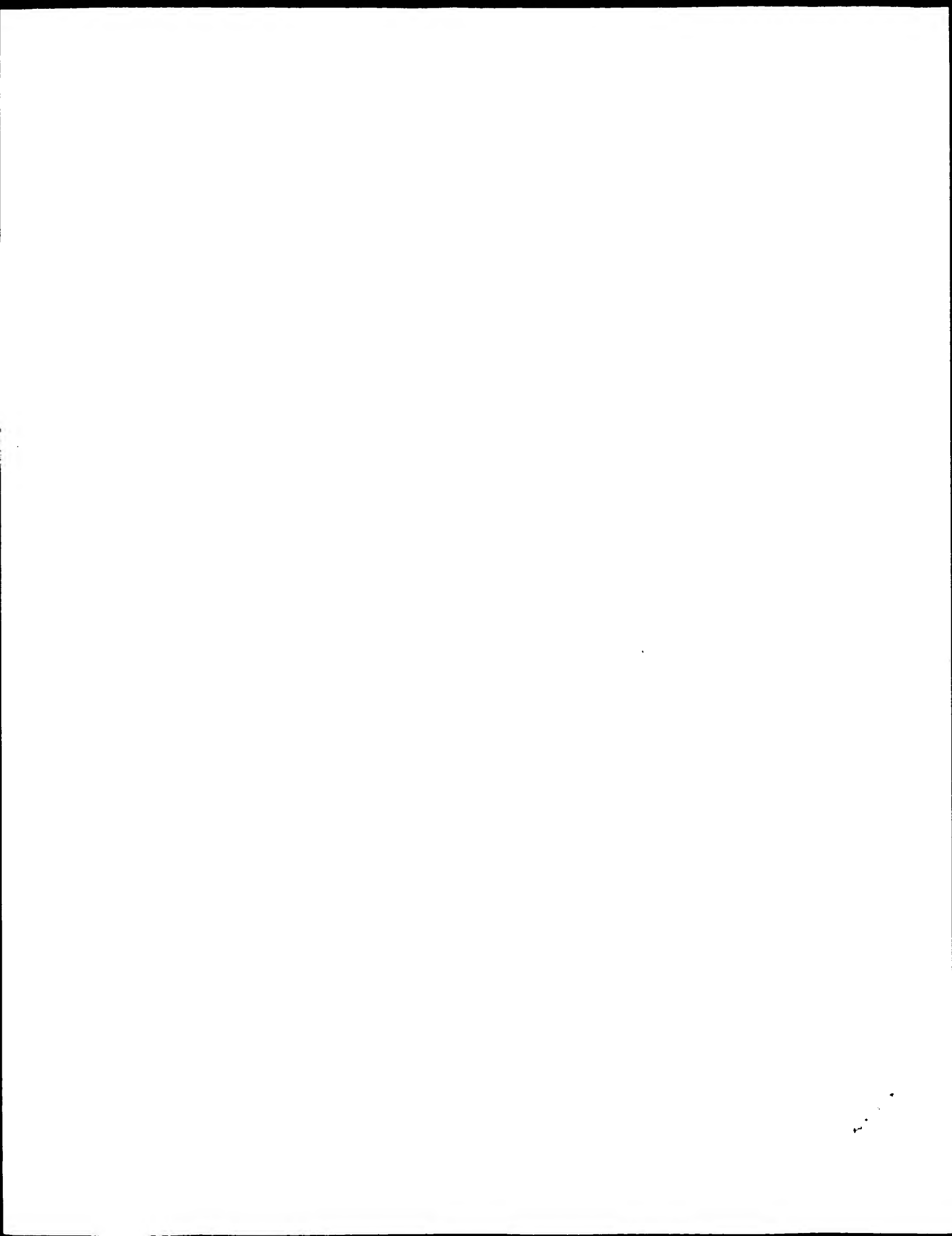
QY 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||
Db 121 VVGGLGYMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
|||||

QY 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||
Db 181 NITIKQHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSMVLFSPPV 240
|||||

QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253
|||||

RESULT 15
S53625
major prion protein - Japanese macaque
C:Species: Macaca fuscata (Japanese macaque)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 13-Aug-1999
C:Accession: S53625; S71053
R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.
J. Mol. Biol. 245, 362-374, 1995
A:Title: Prion protein gene variation among primates.
A:Reference number: S53614; MUID:95139066; PMID:7837269
A:Accession: S53625
A:Molecule type: DNA
A:Residues: 1-253 <SCH>

Search completed: January 28, 2003, 14:20:00
Job time : 7.63425 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 3.86998 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCWMLVLFVATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1435	100.0	253	1	P40156 homo sapien
2	1432	99.8	253	1	P40252 gorilla gor
3	1427	99.4	253	1	P40253 pan troglod
4	1414	98.5	253	1	P40256 pongo pygma
5	1409	98.2	253	1	P40251 colobus gue
6	1405	97.9	253	1	P40257 presbytis f
7	1401	97.6	253	1	P40254 macaca fasc
8	1391.5	97.0	260	1	P40258 salmuri sci
9	1385.5	96.6	252	1	P40247 callithrix
10	1375.5	95.9	252	1	P40249 cebus apell
11	1371.5	95.6	252	1	P51446 ateles pani
12	1356	94.5	246	1	Q95172 cercopithec
13	1353	94.3	246	1	Q95174 cercopithec
14	1352	94.2	246	1	Q95176 cercopithec
15	1331.5	92.8	264	1	P40248 callicebus
16	1331.5	92.8	264	1	P10279 bos taurus
17	1329.5	92.6	254	1	Q950t3 sigmodon hi
18	1329	92.6	241	1	P40255 mandrillus
19	1329	92.6	245	1	P40250 cercopithec
20	1323.5	92.2	254	1	Q60506 cricetus
21	1323.5	92.2	256	1	Q10180 bos taurus
22	1319.5	92.0	254	1	Q60468 cricetus
23	1314.5	91.6	256	1	P23907 ovis aries
24	1312.5	91.5	256	1	P79142 cervus elap
25	1312.5	91.4	256	1	P47852 odocolleus
26	1311.5	91.4	256	1	P52113 capra hircu
27	1311.5	91.4	264	1	P40242 tragelaphus
28	1308.5	91.2	254	1	P40273 mesacricetu
29	1305.5	91.0	259	1	P40245 aotus trivi
30	1305	90.9	237	1	P49927 sus scrofa
31	1301.5	90.7	254	1	P13852 rattus norv
32	1299	90.5	257	1	P40244 mustela vis
33	1298	90.5	255	1	P79141 camelus dro

34	1296.5	90.3	254	1	P40156
35	1290.5	89.9	256	1	P40243 mus musculus
36	1289	89.8	257	1	P40243 tragelaphus
37	1287	89.7	238	1	P52114 mustela put
38	1280.5	89.2	252	1	Q95145 cercocebus
39	1280	89.2	252	1	Q95211 oryctolagus
40	1276.5	89.0	238	1	Q95270 theropithec
41	1271	88.6	255	1	O18754 felis silve
42	1266.5	88.3	232	1	O46501 canis fami
43	1030	71.8	259	1	P40246 ateles geof
44	430.5	30.0	273	1	P51780 trichosurus
45	173	12.1	465	1	P27177 gallus gall
					P10496 phaseolus v

ALIGNMENTS

RESULT 1

ID	PRIO_HUMAN	STANDARD;	PRT;	253 AA.
AC	P04156;			
DT	01-NOV-1986 (Rel. 03, Created)			
DT	01-NOV-1986 (Rel. 03, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (ASCR)			
DE	(CD230 antigen).			
GN	PRNP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86300093; PubMed=3755672;			
RA	Kretschmar H.A., Stowring L.E., Westaway D., Stubblebine W.H.,			
RT	Prusiner S.B., Dearmond S.J.;			
RL	"Molecular cloning of a human prion protein cDNA.";			
	DNA 5:315-324(1986).			
	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvasalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McLay K., McKerray A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.D., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			
RA	Rogers J.;			
RL	"The DNA sequence and comparative analysis of human chromosome 20.";			
	Nature 414:865-871(2001).			
	[3]			
RP	SEQUENCE OF 8-253 FROM N.A.			
RX	MEDLINE=86261776; PubMed=3014653;			
RA	Liao Y.-C.J., Lebo R.V., Clawson G.A., Smuckler E.A.;			
RT	"Human prion protein cDNA: molecular cloning, chromosomal mapping,			

RT and biological implications.";
 RN Science 233:364-367(1986).
 RA [4]
 RP SEQUENCE OF 58-85 AND 111-150 (VARIANT AMYLOID GSS).
 RX MEDLINE=91160504; PubMed=1672107;
 RA Tagliavini F., Prelli F., Ghiso J., Bugiani O., Serban D.,
 RA Prusiner S.B., Farlow M.R., Ghetti B., Frangione B.;
 RT "Amyloid protein of Gerstmann-Straussler-Scheinker disease (Indiana
 RT kindred) is an 11 kd fragment of prion protein with an N-terminal
 RT glycine at codon 58.";
 RL EMBO J. 10:513-519(1991).
 RN [5]
 RP STRUCTURE BY NMR OF 23-230.
 RX MEDLINE=20087216; PubMed=10618385;
 RA Zahn R., Liu A., Luhrs T., Riek R., von Schroetter C., Wuehrich K.;
 RA Lopez Garcia F., Billeter M., Calzolari L., Wider G., Wuehrich K.;
 RT "NMR solution structure of the human prion protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:145-150(2000).
 RN [6]
 RP STRUCTURE BY NMR OF 118-221.
 RX MEDLINE=20359708; PubMed=10900000;
 RA Calzolari L., Lysek D.A., Guntert P., von Schroetter C., Riek R.,
 RA Zahn R., Wuehrich K.;
 RT "NMR structures of three single-residue variants of the human prion
 RT protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:8340-8345(2000).
 RN [7]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=93372867; PubMed=8364585;
 RA Palmer M.S., Collinge J.;
 RT "Mutations and polymorphisms in the prion protein gene.";
 RL Hum. Mutat. 2:168-173(1993).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=94029646; PubMed=8105771;
 RA Prusiner S.B.;
 RT "Genetic and infectious prion diseases.";
 RL Arch. Neurol. 50:1129-1153(1993).
 RN [9]
 RP VARIANT GSS LEU-102.
 RX MEDLINE=89159432; PubMed=2564168;
 RA Hsiao K., Baker H.F., Crow T.J., Poulter M., Owen F.,
 RA Trillinger J.D., Westaway D., Ott J., Prusiner S.B.;
 RT "Linkage of a prion protein missense variant to Gerstmann-Straussler
 RT syndrome.";
 RL Nature 338:342-345(1989).
 RN [10]
 RP VARIANTS LEU-102; VAL-117 AND VAL-129.
 RX MEDLINE=89392018; PubMed=2783132;
 RA Doh-Ura K., Tateishi J., Sasaki H., Kitamoto T., Sakaki Y.;
 RT "Pro-->Leu change at position 102 of prion protein is the most
 RT common but not the sole mutation related to Gerstmann-Straussler
 RT syndrome.";
 RL Biochem. Biophys. Res. Commun. 163:974-979(1989).
 RN [11]
 RP VARIANT FFI ASN-178.
 RX MEDLINE=92195483; PubMed=1347910;
 RA Medori R., Montagna P., Tritschler H.J., Leblanc A., Cortelli P.,
 RA Tinuper P., Lugaresi E., Gambetti P.;
 RT "Fatal familial insomnia: a second kindred with mutation of prion
 RT protein gene at codon 178.";
 RL Neurology 42:669-670(1992).
 RN [12]
 RP VARIANT CJD ASN-178.
 RX MEDLINE=91124933; PubMed=1671440;
 RA Goldfarb L.G., Haltia M., Brown P., Nieto A., Kovanen J.,
 RA McCombie W.R., Trapp S., Gajdusek D.C.;
 RT "New mutation in scrapie amyloid precursor gene (at codon 178) in
 RT Finnish Creutzfeldt-Jakob kindred.";
 RL Lancet 337:425-425(1991).
 RN [13]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=90855709; PubMed=1975028;
 RA Goldfarb L., Mitrova E., Brown P., Toh B.K., Gajdusek D.C.;
 RT "Mutation in codon 200 of scrapie amyloid protein gene in two clusters
 RT of Creutzfeldt-Jakob disease in Slovakia.";
 RL Lancet 336:514-515(1990).
 RN [14]
 RP VARIANT GSS ARG-217.
 RX MEDLINE=93250977; PubMed=1363810;
 RA Hsiao K., Dlouhy S.R., Farlow M.R., Cass C., da Costa M.,
 RA Conneally P.M., Hodes M.E., Ghetti B., Prusiner S.B.;
 RT "Mutant prion proteins in Gerstmann-Straussler-Scheinker disease with
 RT neurofibrillary tangles.";
 RL Nat. Genet. 1:68-71(1992).
 RN [15]
 RP VARIANTS CJD ILE-180 AND ARG-232.
 RX MEDLINE=93213314; PubMed=8461023;
 RA Kitamoto T., Ohta M., Doh-Ura K., Hitoshi S., Terao Y., Tateishi J.;
 RT "Novel missense variants of prion protein in Creutzfeldt-Jakob
 RT disease or Gerstmann-Straussler syndrome.";
 RL Biochem. Biophys. Res. Commun. 191:709-714(1993).
 RN [16]
 RP VARIANT CJD ILE-210.
 RX MEDLINE=94071412; PubMed=7902693;
 RA Pochiarri M., Salvatore M., Cutruzzola F., Genuardi M.,
 RA Alcatelli C.T., Masullo C., Macchi G., Alema G., Gaigani S., Xi Y.G.,
 RA Petraroli R., Silvestrini M.C., Brunori M.;
 RT "A new point mutation of the prion protein gene in Creutzfeldt-Jakob
 RT disease.";
 RL Ann. Neurol. 34:802-807(1993).
 RN [17]
 RP VARIANT GSS LEU-105.
 RX MEDLINE=94077414; PubMed=7902972;
 RA Yamada M., Itoh Y., Fujigasaki H., Naruse S., Kaneko K., Kitamoto T.,
 RA Tateishi J., Otomo E., Hayakawa M., Tanaka J., Matsushita M.,
 RA Miyatake T.;
 RT "A missense mutation at codon 105 with codon 129 polymorphism of the
 RT prion protein gene in a new variant of Gerstmann-Straussler-Scheinker
 RT disease.";
 RL Neurology 43:2723-2724(1993).
 RN [18]
 RP VARIANT GSS LEU-105.
 RX MEDLINE=95213742; PubMed=7699395;
 RA Itoh Y., Yamada M., Hayakawa M., Shozawa T., Tanaka J., Matsushita M.,
 RA Kitamoto T., Tateishi J., Otomo E.;
 RT "A variant of Gerstmann-Straussler-Scheinker disease carrying codon
 RT 105 mutation with codon 129 polymorphism of the prion protein gene: a
 RT clinicopathological study.";
 RL J. Neurol. Sci. 127:77-86(1994).
 RN [19]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=94142912; PubMed=7906019;
 RA Inoue I., Kitamoto T., Doh-Ura K., Shii H., Goto I., Tateishi J.;
 RT "Japanese family with Creutzfeldt-Jakob disease with codon 200 point
 RT mutation of the prion protein gene.";
 RL Neurology 44:299-301(1994).
 RN [20]
 RP VARIANT CJD LYS-200.
 RX MEDLINE=94316708; PubMed=7913755;
 RA Gabizon R., Rosenman H., Meiner Z., Kahana I., Kahana E., Shugart Y.,
 RA Ott J., Prusiner S.B.;
 RT "Mutation in codon 200 and polymorphism in codon 129 of the prion
 RT protein gene in Libyan Jews with Creutzfeldt-Jakob disease.";
 RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 343:385-390(1994).
 RN [21]
 RP VARIANT GSS LEU-102.
 RX MEDLINE=95303274; PubMed=7783876;
 RA Young K., Jones C.K., Piccardo P., Lazzarini A., Golbe L.I.,
 RA Zimmerman T.R., Dickson D.W., McLachlan D.C., St George-Hyslop P.H.,
 RA Lennox A.;
 RT "Gerstmann-Straussler-Scheinker disease with mutation at codon 102
 RT and methionine at codon 129 of PRNP in previously unreported
 RT patients.";
 RL Neurology 45:1127-1134(1995).
 RN [22]


```

Query Match      100.0%; Score 1435; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 6.6e-103;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSPGNGRYPYPPQGGGGWGQP 60
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSPGNGRYPYPPQGGGGWGQP 60

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKSPKPTNNKHMAGAAAAGA 120
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKSPKPTNNKHMAGAAAAGA 120

QY 121 VVGGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMHRYPNQVYRPMDEYSNQNPFVHDCV 180
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 VVGGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMHRYPNQVYRPMDEYSNQNPFVHDCV 180

QY 181 NITIKQHTVTTTKGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFPSPV 240
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 NITIKQHTVTTTKGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFPSPV 240

QY 241 ILLISFLIFLIVG 253
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ILLISFLIFLIVG 253

RESULT 2
PRIO_GORGO STANDARD; PRT; 253 AA.
ID PRIO_GORGO
AC P40252; Q28419;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Gorilla gorilla gorilla (Lowland gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95083661; PubMed=7991600;
RA Cervenakova L., Brown P., Goldfarb L.G., Nagle J., Pettrone K.,
RA Rubenstein R., Dubnick M., Gibbs C.J., Gajdusek D.C.;
RT "Infectious amyloid precursor gene sequences in primates used for
RT experimental transmission of human spongiform encephalopathy.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12159-12162(1994).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; U08300; AAC50089.1; -
DR EMBL; U15166; AAA68633.1; -
DR HSSP; P04156; IQLZ.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 22
  CHAIN 23 230
  PROPEP 231 253
  LIPID 230 230
  DISULFID 179 214
  CARBOHYD 181 181
  CARBOHYD 197 197
  DOMAIN 51 91
  REPEAT 51 59
  REPEAT 60 67
  REPEAT 68 75
  REPEAT 76 83
  REPEAT 84 91
  REPEAT 84 91
  CONFLICT 6 6
  SEQUENCE 253 AA; 27660 MW; E28F4C3FAABCA49E CRC64;

Query Match      99.8%; Score 1432; DB 1; Length 253;
Best Local Similarity 99.6%; Pred. No. 1.1e-102;
Matches 252; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSPGNGRYPYPPQGGGGWGQP 60
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MANLGCMMLVLFVATWSDLGLCKKRPKPGGNTGGSRYPGQSPGNGRYPYPPQGGGGWGQP 60

QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKSPKPTNNKHMAGAAAAGA 120
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQNKKSPKPTNNKHMAGAAAAGA 120

QY 121 VVGGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMHRYPNQVYRPMDEYSNQNPFVHDCV 180
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 VVGGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMHRYPNQVYRPMDEYSNQNPFVHDCV 180

QY 181 NITIKQHTVTTTKGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFPSPV 240
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 NITIKQHTVTTTKGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFPSPV 240

QY 241 ILLISFLIFLIVG 253
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 ILLISFLIFLIVG 253

RESULT 3
PRIO_PANTR
ID PRIO_PANTR STANDARD; PRT; 253 AA.
AC P40253;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Pan troglodytes (Chimpanzee),
OS Hylobates lar (Common gibbon), and
OS Hylobates syndactylus (Siamese), (Symphalangus syndactylus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598, 9580, 9590;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";

```


DR	InterPro: IPR000817; Prion.
DR	Pfam: PF00377; prion; 1.
DR	PRINTS: PR00341; PRION.
DR	SMART: SM00157; PRP; 1.
DR	PROSITE: PS00291; PRION_1; 1.
DR	PROSITE: PS00706; PRION_2; 1.
KW	Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT	SIGNAL 1 22
FT	CHAIN 1 22
FT	PROPEP 23 230
FT	LIPID 231 253
FT	DISULFID 179 214
FT	CARBOHYD 181 191
FT	CARBOHYD 197 187
FT	DOMAIN 51 91
FT	REPEAT 51 59
FT	REPEAT 60 67
FT	REPEAT 68 75
FT	REPEAT 76 83
FT	REPEAT 84 91
SQ	SEQUENCE 253 AA; 27626 MW; 14B17477881F5316 CRC64;
Query Match 98.2%; Score 1409; DB 1; Length 253;	
Best Local Similarity 97.2%; Pred. No. 6.4e-101;	
Matches 246; Conservative 7; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MANLGCMWLVLFVATWSDGLGLCKRKPKGWNTGGSPYCGQSPGNGNRYPPQGGGGWGQP 60
DB	1 MANLGCMWLVLFVATWSDGLGLCKRKPKGWNTGGSPYCGQSPGNGNRYPPQGGGGWGQP 60
QY	61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
DB	61 HGGGWGPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
QY	121 VVGGLGYMGLGSAMSRPIIHFGSDYEDRYRHNHRYPNQVYVRPMDEYSNQNNFVHDCV 180
DB	121 VVGGLGYMGLGSAMSRPIIHFGSDYEDRYRHNHRYPNQVYVRPMDEYSNQNNFVHDCV 180
QY	181 NITIKQHTVTTTKGENFTETDVKMVERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
DB	181 NITIKQHTVTTTKGENFTETDVKMVERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY	241 ILLISFLIFLIVG 253
DB	241 ILLISFLIFLIVG 253
RESULT 6	
ID	PRIO_PREFER STANDARD; PRT; 253 AA.
AC	P40257;
DT	01-FEB-1995 (Rel. 31, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN	PRNP.
OS	Presbytis francoisi.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC	Presbytis.
OX	NCBI_TaxID=33549;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=95139066; PubMed=7837269;
RA	Schatz1 H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT	"Prion protein gene variation among primates.";
RL	J. Mol. Biol. 245:362-374(1995).
CC	!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC	HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC	!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC	"RODS".
CC	!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.


```

RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08304; AAC50092.1; -.
CC HSSP; P04925; IAG2.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC PRINTS; PR00341; PRION.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 229
CC FT PROPEP 230 252
CC FT LPID 229 229
CC FT DISULFID 178 213
CC FT CARBOHYD 180 180
CC FT CARBOHYD 196 196
CC FT DOMAIN 51 90
CC REPEAT 51 58
CC REPEAT 59 66
CC REPEAT 67 74
CC REPEAT 75 82
CC REPEAT 83 90
CC SEQUENCE 252 AA; 27639 MW; B2800B60FD5CE664 CRC64;

Query Match 96.6%; Score 1385.5; DB 1; Length 252;
Best Local Similarity 96.0%; Pred. No. 4e-99;
Matches 243; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGWNTPGGSRYPGQSPGNGNRYPPQGGGCGWQP 60
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGWNTPGGSRYPGQSPGNGNRYPPQ-GGGWGP 59

Qy 61 HGGGWPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQ 120
Db 60 HGGGWPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQHGCGWGPQ 119

Qy 121 VVGGLGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPNQVYRPNQVYRPNQVYRPNQVYRPN 180
Db 120 VVGGLGYMLGSAMSRPLIHFGNDYEDRYRENHRYPNQVYRPNQVYRPNQVYRPNQVYRPNQVYRPN 179

Qy 181 NITIKQHTVTTTKGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
Db 180 NITIKQHTVTTTKGNTFTDVKMERVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 239

Qy 241 ILLISFLIFLIVG 253
Db 240 ILLISFLIFLIVG 252

```

```

RESULT 10
PRIO_CEBAP
ID PRIO_CEBAP STANDARD; PRT; 252 AA.
AC P40249;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
OX NCBI_TaxID=9515;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U08295; AAC50084.1; -.
CC HSSP; P04156; IELG.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC PRINTS; PR00341; PRION.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 229
CC FT PROPEP 230 252
CC FT LPID 229 229
CC FT DISULFID 178 213
CC FT CARBOHYD 180 180
CC FT CARBOHYD 196 196
CC FT DOMAIN 51 90
CC REPEAT 51 58
CC REPEAT 59 66
CC REPEAT 67 74
CC REPEAT 75 82
CC REPEAT 83 90
CC SEQUENCE 252 AA; A2DFCA0AD26D7821 CRC64;

Query Match 95.9%; Score 1375.5; DB 1; Length 252;
Best Local Similarity 95.7%; Pred. No. 2.3e-98;
Matches 242; Conservative 8; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGWNTPGGSRYPGQSPGNGNRYPPQGGGCGWQP 60
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGWNTPGGSRYPGQSPGNGNRYPPQ-GGGWGP 59

```


01-Nov-1997 (Rel. 35, Last sequence update)
01-Nov-1997 (Rel. 35, Last annotation update)
Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
PRNP.
Cercarioxys torquatus atys (Red-crowned mangabey) (Sooty mangabey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Cercopithecus.
NCBI_TaxID=9531;
[1]
SEQUENCE FROM N.A.
der Kuyil A.C., Dekker J.T., Goudsmit J.;
Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
-!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
"RODS".
-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
-!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
-!- SIMILARITY: BELONGS TO THE PRION FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U75385; AAB50628.1; -;
HSSP; P04925; IAG2;
InterPro; IPR000817; Prion.
Pfam; PF00377; prion; 1.
SMART; SM00157; PRP; 1.
PROSITE; PS00291; PRION_1; 1.
PROSITE; PS00706; PRION_2; 1.
Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
NON_TER 1 1
SIGNAL <1 15
CHAIN 16 223
PROPEP 224 246
LIPID 223 223
DISULFID 172 207
CARBOHYD 174 174
CARBOHYD 190 190
DOMAIN 44 84
REPEAT 44 52
REPEAT 53 60
REPEAT 61 68
REPEAT 69 76
REPEAT 77 84
NON_TER 77 84
SEQUENCE 246 AA; 26914 MW; F58679CBCE5ADC7 CRC64;
Query Match 94.2%; Score 1352; DB 1; Length 246;
Best Local Similarity 95.9%; Pred. No. 1.4e-96;
Matches 236; Conservative 10; Mismatches 0; Indels 0; Gaps 0;
QY 8 MLVLFVATWSDLGLCKRKPFGWNTGSGSRYPGQSGPGNRYPPQGGGSGWQPHGGGQ 67
Db 1 MLVLFVATWSDLGLCKRKPFGWNTGSGSRYPGQSGPGNRYPPQGGGSGWQPHGGGQ 60
QY 68 PHGGGSGWQPHGGGQPHGGGQGGGTHSQWNKPSKPTNKKHMAGAAAAGVVGGLG 127
Db 61 PHGGGSGWQPHGGGQPHGGGQGGGTHSQWNKPSKPTNKKHMAGAAAAGVVGGLG 120
QY 128 YMLGSAMSRPLIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNNFVHDCVNITIKQ 187
Db 121 YMLGSAMSRPLIHFGNEYEDRYRENMYRYPNQVYRPMDEYSNQNNFVHDCVNITIKQ 180

QY 188 TVTTTTCGENTETDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPVILLISFL 247
Db 181 TVTTTTCGENTETDVKMMERVVQMCITQYERESQAYYQRGSSMWLFSSPPVILLISFL 240
QY 248 IFLVIG 253
Db 241 IFLVIG 246
RESULT 15
PRIO_CALMO STANDARD; PRT; 241 AA.
AC P40248; 1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 32, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
GN PRNP.
OS Callicebus moloch (Dusky titi).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callicebinae;
OC Callicebus.
OX NCBI_TaxID=9523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
"RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U08312; AAC50100.1; -;
HSSP; P04925; IAG2.
InterPro; IPR000817; Prion.
Pfam; PF00377; prion; 1.
SMART; SM00157; PRP; 1.
PROSITE; PS00291; PRION_1; 1.
PROSITE; PS00706; PRION_2; 1.
Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
NON_TER 1 1
SIGNAL <1 15
CHAIN 16 >241
DISULFID 172 207
CARBOHYD 174 174
CARBOHYD 190 190
DOMAIN 44 84
REPEAT 44 52
REPEAT 53 60
REPEAT 61 68
REPEAT 69 76
REPEAT 77 84
NON_TER 241 241
SEQUENCE 241 AA; 26373 MW; C6D2013EE7CAEC93 CRC64;

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1431	99.7	253	4	Q96E70	homo sapien
2	1427	99.4	253	4	Q8TBQ0	homo sapien
3	1423	99.2	253	4	Q5UP19	homo sapien
4	1409	98.2	285	4	Q75942	homo sapien
5	1381	96.2	246	4	Q60489	homo sapien
6	1366	95.2	245	4	P15216	homo sapien
7	1332	92.8	233	4	Q78446	homo sapien
8	1323.5	92.2	254	11	Q920T4	sigmodon fulv
9	1323.5	92.2	264	6	Q9MZU6	antilocapra
10	1322.5	92.2	254	11	Q8VHV6	apodemus sy
11	1319.5	92.0	256	6	Q95M08	budorcas ta
12	1318	91.8	253	11	Q920T5	meriones un
13	1315.5	91.7	256	6	Q46648	capra hircu
14	1315.5	91.7	256	6	Q02841	odocoileus
15	1313.5	91.5	254	6	Q9TSF8	oryctolagus
16	1312.5	91.5	256	6	Q9TV01	capra hircu


```
Best Local Similarity 88.8%; Pred. No. 8.3e-122;
Matches 253; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQG----- 53
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGQGP 60
QY 54 -----GGGQPGHGGGQPGHGGGQPGHGGGQPGHGGGQPGHGG 88
Db 61 HGGGQPGHGGGQPGHGGGQPGHGGGQPGHGGGQPGHGGGQPGHGG 120
QY 89 WQGGGTHSQNNKPSKPTNNKHMAGAAAAGAVVGLGGYMLGSAMSRPIIHFGSDYEDR 148
Db 121 WQGGGTHSQNNKPSKPTNNKHMAGAAAAGAVVGLGGYMLGSAMSRPIIHFGSDYEDR 180
QY 149 YYRENMHRYPNQVYRPMDEYSNQNNEVHDCVNTIKQHTVTTTKGENTETDVKMMER 208
Db 181 YYRENMHRYPNQVYRPMDEYSNQNNEVHDCVNTIKQHTVTTTKGENTETDVKMMER 240
QY 209 VVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFLIFLIVG 253
Db 241 VVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFLIFLIVG 285

RESULT 5
O60489 PRELIMINARY; PRT; 246 AA.
AC O60489;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein variant (Fragment).
GN PRNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98044029; PubMed=9384372;
RA Samalia H.B., Mari J.J., Vallada H.P., Moura R.P., Simpson A.J.,
RA Brentani R.R.;
RT "A prion-linked psychiatric disorder.";
RL Nature 390:241-241(1997).
DR EMBL; AF030575; AAC05365.1; -.
DR HSP; P04156; IQLZ.
DR InterPro; IPR000817; Prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
FT NON_TER 1
SQ SEQUENCE 246 AA; 26826 MW; 1D9B122EA7D1C18C CRC64;

Query Match 96.2%; Score 1381; DB 4; Length 246;
Best Local Similarity 99.2%; Pred. No. 2.6e-119;
Matches 244; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 MLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGQPGHGGGQ 67
Db 1 MLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGQPGHGGGQ 60
QY 68 PHGGGQPGHGGGQPGHGGGQGGGTHSQNNKPSKPTNNKHMAGAAAAGAVVGLGG 127
Db 61 PHGGGQPGHGGGQPGHGGGQGGGTHSQNNKPSKPTNNKHMAGAAAAGAVVGLGG 120
QY 128 YMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEVHDCVNTIKQ 187
Db 121 YVLSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEVHDCVNTIKQ 180
QY 188 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 247
Db 188 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 247
```

```
Db 181 TVTTTGTGENTETDVKMMERVVEQMCITQYERESQAYYQSGSSMVLFSPPVILLISFL 240
QY 248 IFLIVG 253
Db 241 IFLIVG 246

RESULT 6
Q15216 PRELIMINARY; PRT; 245 AA.
AC Q15216; Q15221;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein.
GN PRP OR PRNP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=91328137; PubMed=1678248;
RA Puckett C., Concannon P., Casey C., Hood L.;
RT "Genomic structure of the human prion protein gene.";
RL Am. J. Hum. Genet. 49:320-329(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Lee I.Y., Westaway D., Smit A.F., Wang K., Cooper C., Yao H.,
RA Prusiner S.B., Hood L.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 9-232 FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=93250789; PubMed=1363802;
RA Diedrich J.F., Knopman D.S., List J.F., Olson K., Frey W.H.,
RA Emory C.R., Sung J.H., Haase A.T.;
RT "Deletion in the prion protein gene in a demented patient.";
RL Hum. Mol. Genet. 1:443-444(1992).
RN [4]
RP SEQUENCE OF 41-85 FROM N.A.
RX MEDLINE=96090306; PubMed=7485229;
RA Perry R.T., Go R.C., Harrell L.E., Acton R.T.;
RT "SSCP analysis and sequencing of the human prion protein gene (PRNP)
RT detects two different 24 bp deletions in an atypical Alzheimer's
RT disease family.";
RL Am. J. Med. Genet. 60:12-18(1995).
DR EMBL; X83416; CAA58442.1; -.
DR EMBL; U29185; AAC78725.1; -.
DR EMBL; W81929; AAB59442.1; -.
DR EMBL; S80743; AAB50649.2; -.
DR EMBL; S80732; AAB50648.2; -.
DR HSP; P04156; IQLZ.
DR InterPro; IPR000817; prion.
DR Pfam; PF00377; prion; 1.
DR PRINTS; PR00341; PRION.
DR SMART; SM00157; PRP; 1.
DR PROSITE; PS00291; PRION_1; 1.
DR PROSITE; PS00706; PRION_2; 1.
KW Prion.
SQ SEQUENCE 245 AA; 26884 MW; 6BF26E0FA3F061AD CRC64;

Query Match 95.2%; Score 1366; DB 4; Length 245;
Best Local Similarity 96.8%; Pred. No. 6.3e-118;
Matches 245; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGQGP 60
Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQG----- 53
QY 61 HGGGQPGHGGGQPGHGGGQPGHGGGQPGHGGGQPGHGGGQPGHGG 120
```



```
Best Local Similarity 89.4%; Pred. No. 4.5e-113;
Matches 227; Conservative 20; Mismatches 6; Indels 1; Gaps 1;

QY 1 MANLGCWMLVLFVATWSDLGLCKKRPKPGGNTGSRYPGOGSPGNRYPPQGGGGWQ 59
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MAHLGYWMLLLFVATWSDVGLCKKRPKPGGNTGSRYPGOGSPGNRYPPQGGGGWQ 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 PHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGTSHQWKNPKPKTKNMKHMAGAAAAG 119
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 PHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGTSHQWKNPKPKTKNMKHMAGAAAAG 120
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 120 AVVGGGLGYMLGSMSRPIIHFGSDYEDRYRYRENMHRYPNQVYRPMDEYSNQNMFVHDC 179
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 AVVGGGLGYMLGSMSRPLIHFGNDYEDRYRYRENMYRYPNQVYRPMDEYSNQNMFVHDC 180
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 180 VNITIKOHTVTTTKGENTFTDVKMMERVVEQMCITQYERESQAYQYRGSSMVLFSPP 239
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 VNITVKOHTVTTTKGENTFTDVKMMERVVEQMCITQYERESQAYQYRGSSMVLFSPP 240
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 240 VILLISFLIFLIVG 253
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 VILLISFLIFLIVG 254
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Search completed: January 28, 2003, 14:18:16
Job time : 13.7848 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 14.0978 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLGCWMLVLFATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1435	100.0	253	17 AAR86715	Human prion protei
2	1435	100.0	253	19 AAW69660	Human prion protei
3	1435	100.0	253	20 AAW85901	Human prion protei
4	1435	100.0	253	21 AAB15035	Human prion protei
5	1435	100.0	253	21 AAB06272	Human PrP pr
6	1435	100.0	253	21 AAY81485	Human prion protei
7	1435	100.0	253	22 AAG55853	Human prion protei
8	1435	100.0	253	22 AAB82112	Human PrP. Homo s
9	1435	100.0	253	22 AAB61770	Human prion protei
10	1435	100.0	253	23 ABP51787	Human prion protei

11	1435	100.0	253	23 AAE15603	Human PrP protei.
12	1435	100.0	253	23 ABE04426	Human prion protei
13	1427	99.4	253	20 AAY07994	Human prion protei
14	1425	99.3	253	22 AAB72338	Human prion protei
15	1423	99.2	253	23 AAB79575	Human prion protei
16	1422	99.1	253	23 AAB72341	Gorilla prion prote
17	1417	98.7	253	22 AAB72339	Chimpanzee prion p
18	1417	98.7	253	22 AAB72345	Gibbon prion prote
19	1417	98.7	253	22 AAB72356	Slamang prion prote
20	1404	97.8	253	22 AAB72340	Orangutan prion pr
21	1399	97.5	253	22 AAB72353	Guereza prion prote
22	1395	97.2	253	22 AAB72355	Prion protein cell
23	1392	97.0	253	22 AAB72350	Marmoset prion pro
24	1391	96.9	253	22 AAB72344	Rhesus monkey prio
25	1391	96.9	253	22 AAB72346	Prion protein cell
26	1391	96.9	253	22 AAB72347	Prion protein cell
27	1391	96.9	253	22 AAB72348	Prion protein cell
28	1391	96.9	253	22 AAB72349	Prion protein cell
29	1391	96.9	253	22 AAB72351	Hamadryas prion pr
30	1382	96.3	253	22 AAB72354	Capuchin prion pro
31	1376.5	95.9	260	22 AAB72343	Monkey prion prote
32	1332.5	92.9	264	22 AAB72361	Cow prion protei
33	1331.5	92.8	264	22 AAB82113	Bovine PrP. Bos t
34	1331.5	92.8	264	22 AAB61769	Bovine prion prote
35	1331.5	92.8	264	23 AAE15604	Bovine PrP protei
36	1329	92.6	263	17 AAR86716	Bovine prion prote
37	1329	92.6	263	19 AAW69661	Bovine prion prote
38	1329	92.6	263	20 AAW85902	Bovine prion prote
39	1329	92.6	263	22 AAG65854	Bovine prion prote
40	1329	92.6	263	23 ABP51788	Bovine prion prote
41	1326.5	92.4	264	23 ABB04424	Bovine prion prote
42	1323	92.2	265	23 AAM50889	Bovine prion prote
43	1322.5	92.2	264	20 AAY07995	Bovine prion prote
44	1321	92.1	255	17 AAR86717	Sheep prion protei
45	1321	92.1	255	19 AAW69662	Sheep prion protei

ALIGNMENTS

RESULT 1
AAR86715
ID AAR86715 standard; protein; 253 AA.
XX
AC AAR86715;
XX 15-OCT-1996 (first entry)
DT Human prion protein, HuPrP.
XX
DE
XX
KW Chimeric gene; chimeric prion; transgenic animal; diagnosis;
KW spongiform encephalopathy; PrP; central nervous system; CNS;
KW Creutzfeld-Jakob disease; CJD; BSE.
XX
OS Homo sapiens.
XX
PN WO9531466-A1.
XX
PD 23-NOV-1995.
XX
PF 10-APR-1995; 95WO-US04426.
XX
PR 13-MAY-1994; 94US-0242188.
XX (REGC) UNIV CALIFORNIA.
XX Prusiner SB, Scott MR, Telling G;
XX WPI; 1996-010868/01.
XX Chimeric prion protein gene - for formation of a transgenic animal
XX susceptible to prion infection by prion(s) normally specific for a
XX different species

PA (REGC) UNIV CALIFORNIA.
 XX Burton DR, Prusiner SB, Williamson RA;
 XX WPI; 1999-058996/05.
 XX Antibody specific for scrapie isoform of prion protein - useful for
 PT diagnosis and therapy
 XX
 XX Disclosure; Columns 41-42; 58pp; English.
 XX
 XX This represents a human prion protein (PrP) sequence. The invention
 CC relates to an antibody that is capable of binding to native PrP(Sc), the
 CC scrapie isoform of PrP. The antibody is produced by a method that
 CC comprises synthesising a library of antibodies on phages, contacting the
 CC phages with a composition containing PrP proteins, isolating phages that
 CC bind to native PrP(Sc) in situ, obtaining an antibody from the phages,
 CC and optionally analysing the phages to determine a nucleic acid sequence
 CC encoding an amino acid sequence to which the native PrP(Sc) binds. The
 CC antibody is used to detect disease-associated PrP, especially in
 CC Creutzfeldt-Jakob disease (CJD) and distinguish it from normal PrP. They
 CC can also be used to neutralise the infectivity of PrP(Sc). Assays using
 CC the antibodies can be used to screen for disease-associated PrP in
 CC pharmaceutical products, foods and cosmetics or for therapeutic purposes.
 XX
 XX Sequence 253 AA;
 SQ
 Query Match 100.0%; Score 1435; DB 20; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCMWLVLVATWSDLGLCKRKPGGWNTGSGRYPGQSGPGNRYPPQGGGWWGP 60
 Db 1 MANLGCMWLVLVATWSDLGLCKRKPGGWNTGSGRYPGQSGPGNRYPPQGGGWWGP 60
 QY 61 HGGGWWGPHGGGWWGPHGGGWWGPHGGGWWGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 Db 61 HGGGWWGPHGGGWWGPHGGGWWGPHGGGWWGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMYLFSPPV 240
 Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMYLFSPPV 240
 QY 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253
 RESULT 4
 AAB15035
 ID AAB15035 standard; Protein; 253 AA.
 XX
 XX AAB15035;
 XX 18-DEC-2000 (first entry)
 XX Human prion protein.
 XX
 XX Prion; PrP; guanidine thiocyanate; gdnSCN; TSE; BSE;
 KW transmissible spongiform encephalopathy;
 KW bovine spongiform encephalopathy; sheep; cattle; human.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..22
 FT /note= "N terminal signal peptide"
 FT 23..231
 FT /label= "Mature protein"
 FT

FT Cleavage-site 60..89
 FT /note= "protease sensitive site"
 FT Peptide 232..253
 FT /note= "C terminal signal peptide"
 XX
 PN WO200048003-A1.
 XX 17-AUG-2000.
 XX 09-FEB-2000; 2000WO-NL00079.
 XX
 XX 11-FEB-1999; 99EP-0200391.
 XX (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.
 XX
 PI Garssen GJ, Jacobs JG, Langeveld JPM, Smits MA, Van Keulen LJM;
 PI Schreuder BEC, Bossers A;
 XX WPI; 2000-506099/45.
 DR
 XX Use of guanidine thiocyanate for reducing risk of false-positive
 PT results in testing mammalian sample for aberrant prion protein, useful
 PT for detection of transmissible spongiform encephalopathies -
 XX
 XX Disclosure; Fig 1; 49pp; English.
 XX
 XX The present invention relates to a method for reducing the risk of
 CC scoring a false positive test result in testing a sample for aberrant
 CC prion protein. The method involves the use of guanidine thiocyanate
 CC (gdnSCN) or its functional equivalent. This test is highly useful for
 CC testing for transmissible spongiform encephalopathies (TSEs) such as
 CC BSE (bovine spongiform encephalopathy). The method allows a faster,
 CC simpler and more reliable method for monitoring cattle and sheep for
 CC the presence of aberrant prion protein before it reaches the human
 CC and animal food chain. The present sequence is the human prion protein.
 XX
 XX Sequence 253 AA;
 SQ
 Query Match 100.0%; Score 1435; DB 21; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCMWLVLVATWSDLGLCKRKPGGWNTGSGRYPGQSGPGNRYPPQGGGWWGP 60
 Db 1 MANLGCMWLVLVATWSDLGLCKRKPGGWNTGSGRYPGQSGPGNRYPPQGGGWWGP 60
 QY 61 HGGGWWGPHGGGWWGPHGGGWWGPHGGGWWGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 Db 61 HGGGWWGPHGGGWWGPHGGGWWGPHGGGWWGTHSQWNKPSKPTNKKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 QY 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMYLFSPPV 240
 Db 181 NITIKQHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMYLFSPPV 240
 QY 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253
 RESULT 5
 AAB06272
 ID AAB06272 standard; Protein; 253 AA.
 XX
 XX AAB06272;
 XX 16-OCT-2000 (first entry)
 XX Human PrP prion protein.
 XX

KW Human; PrP; prion; Creutzfeldt-Jakob disease; CJD; neuroprotective;
 KW vaccine; beta-form PrP; Kuru; transmissible mink encephalopathy;
 XX bovine spongiform encephalopathy; BSE.
 OS Homo sapiens.
 XX WO200026238-A2.
 PN 11-MAY-2000.
 PD 02-NOV-1999; 99WO-GB03617.
 PF 04-NOV-1998; 98GB-0024091.
 PR 18-MAR-1999; 99GB-0006217.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA Collinge J, Clarke AR, Jackson GS;
 PI WPI; 2000-365570/31.
 DR N-PSDB; AAA57174.
 XX Novel method of producing beta-form prion proteins, related antibodies,
 PT and binding agents useful in treatment and diagnosis of prion diseases
 PT .
 XX Disclosure; Fig 1; 115pp; English.
 PS The present sequence is the human prion protein encoded by the PrP
 CC gene. A beta-form of the protein has been produced which has more
 CC beta-sheet than alpha-helix structure, can exist as a monomer and
 CC can retain solubility in an aqueous solution in the absence of a
 CC denaturant. The beta-form PrP, a beta-form binding agent or aggregate can
 CC be used in the preparation or manufacture of a composition for the
 CC prevention, treatment and/or diagnosis of a prion disease, e.g. Kuru,
 CC Creutzfeldt-Jakob disease (CJD), transmissible mink encephalopathy,
 CC chronic wasting disease of mule deer and elk, and bovine spongiform
 CC encephalopathy (BSE). By measuring the levels of beta-form in a sample,
 CC agents capable of preventing, reducing and/or reversing the conversion of
 CC a PrP to a beta-form can be identified. The beta-form of the PrP or a
 CC non-fibrillar aggregate can be used as a vaccine against a prion disease.
 CC The beta-form can also be used to diagnose a predisposition to or the
 CC presence of a prion disease by monitoring aggregation.
 XX Sequence 253 AA;
 SQ
 Query Match 100.0%; Score 1435; DB 21; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 DB 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 DB 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 QY 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 DB 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 QY 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 DB 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 DB 241 ILLISFLIFLIVG 253
 RESULT 7
 AAG65853
 ID AAG65853 standard; protein; 253 AA.
 XX
 AC AAG65853;
 XX
 DT 11-FEB-2002 (first entry)

ID AAY81485 standard; Protein; 253 AA.
 XX
 AC AAY81485;
 XX
 DT 03-JUL-2000 (first entry)
 XX
 DE Human prion protein.
 XX
 KW Prion protein; human; antibody; monoclonal; prion-related disease;
 KW Creutzfeldt-Jakob disease; CJD; diagnosis.
 XX
 OS Homo sapiens.
 XX JP2000060551-A.
 PN 29-FEB-2000.
 PD 13-AUG-1998; 98JP-0241145.
 PF 13-AUG-1998; 98JP-0241145.
 PR (SEK) SEIKAKAKU KOGYO CO LTD.
 PA WPI; 2000-249674/22.
 XX Anti-prion antibody specific for a particular amino acid sequence -
 XX useful in diagnosis and research of prion related diseases
 PT
 PS Disclosure; Page 8; 9pp; Japanese.
 XX The invention relates to an anti-prion protein antibody, preferably a
 CC mouse-derived monoclonal antibody, which binds residues 96-114 of human
 CC prion protein (peptide sequence given in AAY81484). The anti-prion
 CC antibody is useful as a research tool and for the diagnosis of
 CC prion-related diseases such as Creutzfeldt-Jakob disease (CJD). The
 CC antibody is highly specific for human prion protein, and is relatively
 CC inexpensive to manufacture. This sequence represents full-length human
 CC prion protein.
 XX Sequence 253 AA;
 SQ
 Query Match 100.0%; Score 1435; DB 21; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 DB 1 MANLGCWMLVLFVATWSDGLCKKRPKPGWNTGSGRYPGQSPGNGRYPPOGGGGWGP 60
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA 120
 QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 DB 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNFVHDCV 180
 QY 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 DB 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSWLVFSSPPV 240
 QY 241 ILLISFLIFLIVG 253
 DB 241 ILLISFLIFLIVG 253
 RESULT 7
 AAG65853
 ID AAG65853 standard; protein; 253 AA.
 XX
 AC AAG65853;
 XX
 DT 11-FEB-2002 (first entry)


```

XX DE Human prion protein (PrP) sequence.
XX DE
XX KW PrP; prion protein; Creutzfeldt-Jakob disease; familial insomnia; PrP-Sc;
XX KW scrapie; Gerstmann-Strassler-Scheinker disease.
XX OS Homo sapiens.
XX PN US6290954-B1.
XX PD 18-SEP-2001.
XX PF 06-MAR-1998; 98US-0036579.
XX PR 13-SEP-1996; 96US-0713939.
XX PR 14-SEP-1995; 95US-0528104.
XX PA (SCRI ) SCRIPPS RES INST.
XX PI Prusiner SB, Williamson RA, Burton DR;
XX WPI; 2001-637939/73.
XX
XX Detecting a scrapie isoform of the prion protein (PrP-Sc) in a source,
XX particularly useful for detecting e.g. Creutzfeldt-Jakob disease or
XX Gerstmann-Strassler-Scheinker disease, by contacting the source with
XX PrP-Sc antibodies -
XX
XX Disclosure; Fig 2; 58pp; English.
XX
XX The invention provides a method for detecting a scrapie isoform of the
XX prion protein (PrP-Sc) in a source. The method involves contacting the
XX source suspected of containing native PrP-Sc with a diagnostic amount of
XX an antibody characterized by its ability to bind to native PrP-Sc in
XX situ. The method is useful for detecting PrP-Sc in a source, which is
XX particularly useful for detecting Creutzfeldt-Jakob disease, fatal
XX familial insomnia or Gerstmann-Strassler-Scheinker disease. The present
XX sequence represents the human PrP sequence.
XX
XX Sequence 253 AA;
XX
XX Query Match 100.0%; Score 1435; DB 22; Length 253;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-135;
XX Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGGSRYPGQSPGNGRYPPOGGGGWGQP 60
DB 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGGSRYPGQSPGNGRYPPOGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
QY 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
DB 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSWVLFSSPPV 240
DB 181 NITIKOHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSWVLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
DB 241 ILLISFLIFLIVG 253
XX
XX RESULT 8
XX ID AAB82112
XX AC AAB82112 standard; Protein; 253 AA.
XX DT 29-JUN-2001 (first entry)

```

```

XX DE Human Prp.
XX DE
XX KW Human; Prp; cerebroprotective; PrP conversion inhibitor; prion protein;
XX KW transmissible spongiform encephalopathy; TSE; neurodegenerative disease;
XX KW protease-sensitive prion protein; PrPsen;
XX KW protease-resistant prion protein; PrPres.
XX OS Homo sapiens.
XX PN US6211149-B1.
XX PD 03-APR-2001.
XX PF 03-AUG-1998; 98US-0128450.
XX PR 03-AUG-1998; 98US-0128450.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Chesebro BW, Caughey BW, Chabry J, Priola S;
XX WPI; 2001-315407/33.
XX
XX New peptide comprises hamster prion protein fragment that inhibits
XX conversion of prion protein from protease-sensitive to
XX protease-resistant form, useful for diagnosis and treatment of
XX spongiform encephalopathies -
XX
XX Disclosure; Column 33-34; 31pp; English.
XX
XX Transmissible spongiform encephalopathies (TSE) are fatal
XX neurodegenerative diseases. These diseases are characterized by the
XX formation and accumulation, in the brain, of an abnormal proteinase K
XX resistant isoform (PrPres) of a normal protease-sensitive host-encoded
XX prion protein (PrPsen). The present invention relates to peptides
XX comprising a hamster, human or murine prion protein (PrP) fragment which
XX specifically inhibit the conversion of protease-sensitive prion protein
XX (PrPsen) to protease-resistant prion protein (PrPres). The present
XX sequence is the protein sequence for human PrP. The peptides of the
XX present invention are useful for diagnosis and treatment of TSE
XX diseases.
XX
XX Sequence 253 AA;
XX
XX Query Match 100.0%; Score 1435; DB 22; Length 253;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-135;
XX Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGGSRYPGQSPGNGRYPPOGGGGWGQP 60
DB 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGGSRYPGQSPGNGRYPPOGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNNKHMAGAAAAGA 120
QY 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
DB 121 VVGLGGLYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSWVLFSSPPV 240
DB 181 NITIKOHTVTTTNGENFTETDVKMERVVQMCITQYERESQAYYQRGSSWVLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
DB 241 ILLISFLIFLIVG 253
XX
XX RESULT 9
XX ID AAB61770
XX ID AAB61770 standard; protein; 253 AA.

```


Db	181	NIYKQHTVTTTKGNGFTETDVKMHWRYEQMCITQYRESQAYQYRGSMVLFSSPPV	240
QY	241	ILLISFLIFLIVG	253
Db	241	ILLISFLIFLIVG	253
RESULT 10			
ABP51787			
ID	ABP51787	standard; protein; 253 AA.	
XX	AC	ABP51787:	
XX	DT	03-OCT-2002 (first entry)	
XX	XX	Human prion protein (PrP) SEQ ID NO:2.	
XX	XX	Prion protein; PrP; scrapie; PrPSc; prion disease; immunoassay; detection.	
KW	XX	Homo sapiens.	
OS	XX	US6372214-B1.	
XX	PN	16-APR-2002.	
XX	PD	13-APR-2000; 2000US-0550374.	
XX	PF	13-SEP-1996; 96US-0713939.	
XX	PR	06-MAR-1998; 98US-0036579.	
XX	PR	14-SEP-1995; 95US-0528104.	
XX	XX	(REGC) UNIV CALIFORNIA.	
PA	XX	(SCRI) SCRIPPS RES INST.	
XX	PI	Prusiner SB, Williamson RA, Burton DR;	
XX	PI	WPI; 2002-433675/46.	
XX	DR		
XX	PT	Immunoassays for detecting scrapie isoforms of prion protein (PrPSc)	
XX	PT	and for purifying PrPSc from samples, useful e.g. in diagnosing PrPSc	
XX	PT	disease and testing pharmaceuticals for contamination -	
XX	PS	Disclosure; Fig 2; 58pp; English.	
XX	XX	The present invention describes methods for detecting scrapie isoforms	
XX	CC	of prion protein (PrPSc) infection in dead animals, purifying materials	
XX	CC	suspected of containing PrPSc proteins and treating materials, using	
XX	CC	antibodies specific for PrPSc. Also described: (1) method of determining	
XX	CC	PrPSc infection in a dead animal, comprising: (a) extracting tissue from	
XX	CC	an animal that has died; (b) contacting the tissue with an antibody	
XX	CC	characterised by its ability to bind to native PrPSc in situ (the	
XX	CC	antibody binds to a form of PrPSc specific to the animal that has died);	
XX	CC	and (c) determining if the antibody has bound to PrPSc (the presence of	
XX	CC	PrPSc in the tissue is indicative of PrPSc infection); (2) a method of	
XX	CC	purifying a material suspected of containing a PrPSc protein, comprising:	
XX	CC	(a) contacting the material with an antibody (characterized by its	
XX	CC	ability to bind native PrPSc in situ) which is bound to a support	
XX	CC	surface; and (b) removing material not bound to the antibody; (3) a	
XX	CC	method of treating a material, comprising applying (to the material) an	
XX	CC	antibody that binds native PrPSc in situ. The methods are used for	
XX	CC	diagnosing and detecting prion disease (scrapie) in dead animal tissue	
XX	CC	(i.e. immunoassays), for separating PrPSc proteins from biological	
XX	CC	samples (i.e. immunopurification) and for treating materials. The present	
XX	CC	sequence represents the human prion protein (PrP) which is given in the	
XX	CC	embodiment of the present invention.	
XX	XX	Sequence 253 AA:	
Query Match		100.0%; Score 1435; DB 23; Length 253;	
Best Local Similarity		100.0%; Pred. No. 2.4e-135;	
Matches 253; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	

QY 1 MANLGCMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 60
Db 1 MANLGCMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQ 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQ 120
QY 121 VVGLGGMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 180
Db 121 VVGLGGMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 180
QY 181 NITIKOHTVTTTGGKGFETDVKMERVVEQMCITQYERESQAYYQYRGSSWVLFSSPPV 240
Db 181 NITIKOHTVTTTGGKGFETDVKMERVVEQMCITQYERESQAYYQYRGSSWVLFSSPPV 240
QY 241 ILLISFLIFLI 253
Db 241 ILLISFLIFLI 253
RESULT 11
AAE15603
ID AAE15603 standard; Protein: 253 AA.
AC AAE15603;
DT 12-MAR-2002 (first entry)
DE Human PrP protein.
KW Protease resistant prion protein; PrPres; Creutzfeldt-Jakob disease;
KW protease sensitive prion protein; PrPsen; therapy; neuroprotectant;
KW transmissible spongiform encephalopathy; fatal familial insomnia;
KW Gerstmann-Strausler-Scheinker syndrome; scrapie; kuru; human.
OS Homo sapiens.
PN US2001041790-A1.
XX 15-NOV-2001.
XX 30-MAR-2001; 2001US-0823494.
XX 12-MAY-1998; 98US-085160P.
XX 03-AUG-1998; 98US-0128450.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Chesebro BW, Caughey BW, Chabry J, Priola S;
XX WPI; 2002-065984/09.
XX Peptides which can inhibit conversion of prion proteins from protease
XX sensitive to protease resistant forms, are useful in treating
XX transmissible spongiform encephalopathies -
XX Claim 45; Page 18-19; 33pp; English.
XX The present invention relates to peptides comprising a peptide region of
XX prion protein (PrP) and which are capable of inhibiting conversion of
XX protease sensitive prion protein (PrPsen) to protease resistant prion
XX protein (PrPres). The peptides are useful for inhibiting formation of
XX protease resistant prion proteins, such as those associated with
XX transmissible spongiform encephalopathies, e.g., Creutzfeldt-Jakob
XX disease, kuru, Gerstmann-Strausler-Scheinker syndrome, fatal familial
XX insomnia or scrapie. The peptides can be used as diagnostic agents, e.g.,
XX to detect the presence of PrPres in body fluids such as blood. They may
XX also be used to treat or prevent diseases such as the above. They may
XX also be used to design analogues, derivatives or mimetics for use as
XX inhibitors of conversion of PrPsen to PrPres. The present sequence is
XX human PrP protein.

SQ Sequence 253 AA;
Query Match 100.0%; Score 1435; DB 23; Length 253;
Best Local Similarity 100.0%; Pred. No. 2,4e-135;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANLGCMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 60
Db 1 MANLGCMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQ 120
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQ 120
QY 121 VVGLGGMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 180
Db 121 VVGLGGMWLVFVATWSDLGLCKKRPKPGWNTGGSRYPGQSPGNGRYPPOGGGGWGQ 180
QY 181 NITIKOHTVTTTGGKGFETDVKMERVVEQMCITQYERESQAYYQYRGSSWVLFSSPPV 240
Db 181 NITIKOHTVTTTGGKGFETDVKMERVVEQMCITQYERESQAYYQYRGSSWVLFSSPPV 240
QY 241 ILLISFLIFLI 253
Db 241 ILLISFLIFLI 253
RESULT 12
ABB04426
ID ABB04426 standard; Protein: 253 AA.
XX ABB04426;
XX 04-MAR-2002 (first entry)
XX Human prion protein PrP.
XX Human; prion protein; PrP; antiviral; HIV; prion disease; kuru; virucide;
KW antibacterial; neuroprotective; anti-HIV; Creutzfeldt-Jakob disease;
KW Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;
KW bovine spongiform encephalitis; scrapie.
OS Homo sapiens.
XX WO200183747-A2.
XX 08-NOV-2001.
XX 30-APR-2001; 2001WO-FR01336.
XX 28-APR-2000; 2000FR-0005535.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Leblanc P, Darlix J, Gabus-Darlix C;
XX WPI; 2002-049350/06.
XX New polypeptides, useful as antiviral agents, comprise their prion
XX proteins able to bind nucleic acid, nucleocapsid proteins, and ligands
XX for use as antiprion agents -
XX Disclosure; Fig 9; 80pp; French.
XX The present invention relates to normal (PrPc) or abnormal (PrPsc) human
XX or animal prion proteins which are able to bind to DNA or RNA,
XX particularly of viral, especially retroviral, origin and to nucleocapsid
XX proteins (NCP) of human or animal retroviruses. These can be used as
XX antiviral agents, particularly against human immune deficiency virus
XX (HIV), and in the treatment of prion diseases including Creutzfeldt-Jakob
XX disease, Gerstmann-Strausler-Scheinker disease, fatal familial
XX insomnia, kuru, bovine spongiform encephalitis and scrapie. The present
XX sequence is the human PrP protein sequence.

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 1435; DB 23; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MANLGCMVLVLFVATWSDLGCKKRPKPGGWNTPGGSRYPGQSPGGRNRYPPQGGGGWGP 60
 Db 1 MANLGCMVLVLFVATWSDLGCKKRPKPGGWNTPGGSRYPGQSPGGRNRYPPQGGGGWGP 60

Qy 61 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNMKHMAGAAAAGA 120
 Db 61 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNMKHMAGAAAAGA 120

Qy 121 VVGLGGYMLGMSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
 Db 121 VVGLGGYMLGMSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180

Qy 181 NITIKQHTVTTTNGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 Db 181 NITIKQHTVTTTNGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240

Qy 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253

RESULT 13
 AAY07994
 ID AAY07994 standard; protein; 253 AA.
 XX AC AAY07994;
 XX DT 08-JUL-1999 (first entry)
 XX DE Human prion protein.
 XX KW Prion protein; PrP; human; polyclonal antiserum; immunoassay;
 XX KW detection; bovine; murine.
 XX OS Homo sapiens.
 XX PN DE19745443-A1.
 XX PD 22-APR-1999.
 XX PF 15-OCT-1997; 97DE-1045443.
 XX PR 15-OCT-1997; 97DE-1045443.
 XX PA (HERZ/) HERZOG-MESMER A.
 XX PI Kiselev OI, Mesmer AH, Scheller A;
 XX PS WPI; 1999-255775/22.
 XX PT Diagnostic polyclonal antiserum specific for prion protein -
 XX PT obtained by immunisation with metal-containing polypeptide
 XX PS Claim 3; Fig 1; 12pp; German.
 XX CC This invention describes a novel process for producing a polyclonal
 CC antiserum against a human or animal prion protein (PrP) which can be
 CC used in immunoassays for detecting PrP's. The method comprises (a)
 CC selecting a polypeptide that has a length of at least 10 amino acids and
 CC has an amino acid sequence at least 70% homologous to that of human,
 CC bovine or murine PrP in a region of at least 10 consecutive amino acids
 CC (b) binding a metal to the polypeptide by reaction with a metal compound
 CC and (c) injecting the metal-containing polypeptide into a host animal,
 CC optionally together with adjuvants, to induce production of a polyclonal
 CC antiserum.

SQ Sequence 253 AA;

Query Match 99.4%; Score 1427; DB 20; Length 253;
 Best Local Similarity 99.6%; Pred. No. 1.5e-134;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MANLGCMVLVLFVATWSDLGCKKRPKPGGWNTPGGSRYPGQSPGGRNRYPPQGGGGWGP 60
 Db 1 MANLGCMVLVLFVATWSDLGCKKRPKPGGWNTPGGSRYPGQSPGGRNRYPPQGGGGWGP 60

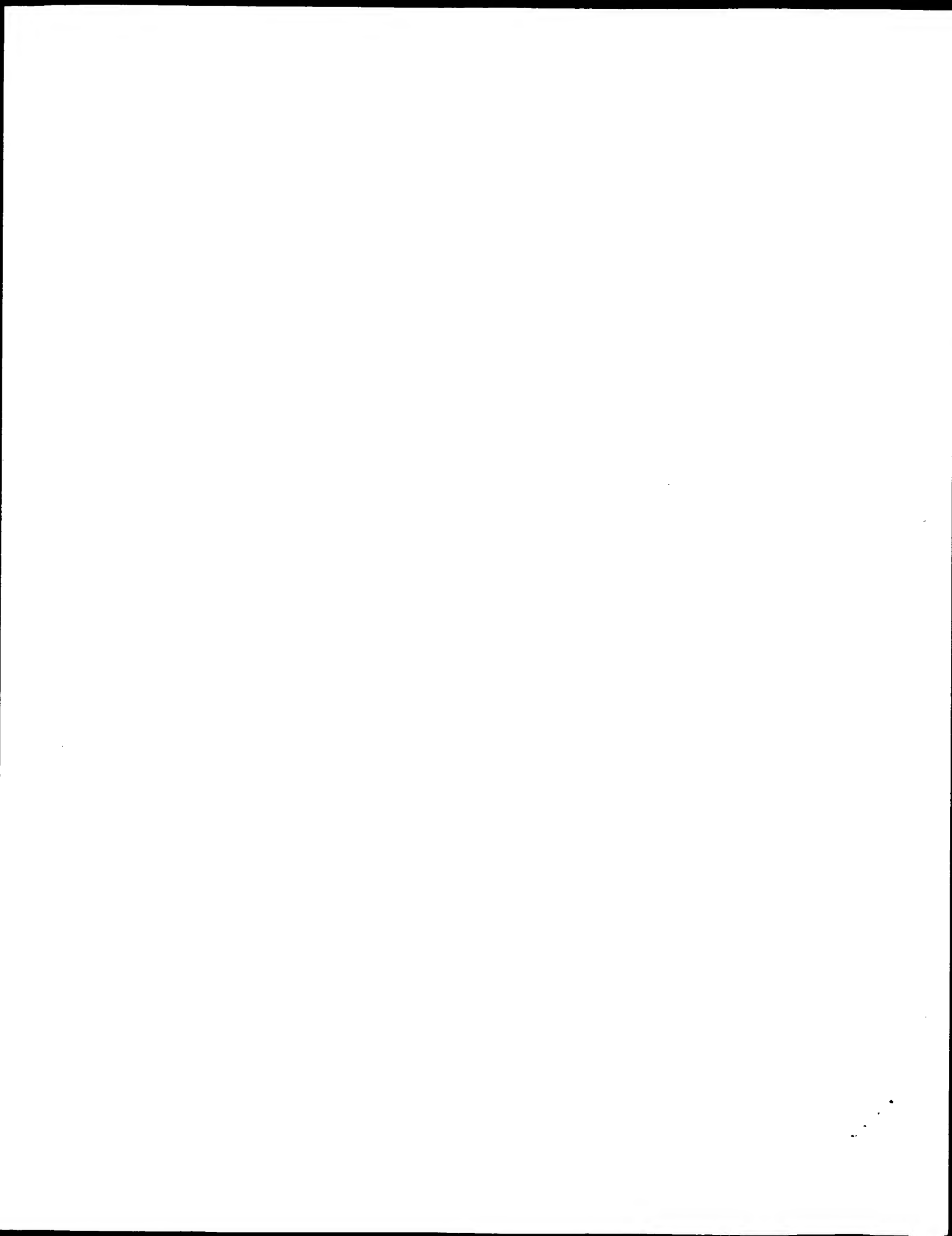
Qy 61 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNMKHMAGAAAAGA 120
 Db 61 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNMKHMAGAAAAGA 120

Qy 121 VVGLGGYMLGMSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180
 Db 121 VVGLGGYMLGMSAMSRPIIHFGSDYEDRYRENMHRYPNQVYRPMDEYSNQNMFVHDCV 180

Qy 181 NITIKQHTVTTTNGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 Db 181 NITIKQHTVTTTNGENFTETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240

Qy 241 ILLISFLIFLIVG 253
 Db 241 ILLISFLIFLIVG 253

RESULT 14
 AAB72338
 ID AAB72338 standard; Peptide; 253 AA.
 XX AC AAB72338;
 XX DT 17-MAY-2001 (first entry)
 XX DE Human prion protein cellular form (PrPc) amino acid sequence.
 XX KW Prion protein; cellular form; PrPc; stable region; antibody; BSE; CJD;
 XX KW prion disease; spongiform encephalopathies; Scrapie; human;
 XX KW bovine spongiform encephalopathy; BSE; Creutzfeldt-Jakob disease.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT Region 176..221
 XX FT /note= "Stable region, specifically claimed in claim 3"
 XX PN WO200107479-A2.
 XX PD 01-FEB-2001.
 XX PF 25-JUL-2000; 2000WO-GB02873.
 XX PR 27-JUL-1999; 99GB-0017491.
 XX PR 30-JUL-1999; 99GB-0017878.
 XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX PI Collinge J, Clarke AR, Waltho JP, Jackson GS, Hosszu LLP;
 XX PS WPI; 2001-168538/17.
 XX CC New prion peptide for treating, preventing and/or diagnosing prion
 CC diseases e.g. scrapie in sheep, bovine spongiform encephalopathies in
 CC cows and Creutzfeldt-Jakob disease in humans
 CC Claim 3; Fig 5; 69pp; English.
 CC This invention relates to a peptide fragment of a cellular form of prion
 CC protein PrPc located around a disulphide bond found in PrPc. The stable
 CC structure is a specific marker of PrPc but not soluble prion protein
 CC (PrPsc). The PrPc peptide sequences can be used to generate an antibody
 CC or binding agent that binds PrPc. The antibody is used to detect or



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 3.31713 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-3
Perfect score: 1435
Sequence: 1 MANUGCWLVLVATNSDLG.....LFSSPVILLISFLIFLVG 253

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #	Query			DB	ID	Description
		Score	Match	Length			
1	1435	100.0	253	10	US-09-904-987-3	Sequence 3, Appli	
2	1435	100.0	253	10	US-09-943-906-2	Sequence 2, Appli	
3	1431	99.7	253	10	US-09-919-172-57	Sequence 57, Appl	
4	1329	92.6	263	10	US-09-943-906-3	Sequence 3, Appli	
5	1326.5	92.4	256	9	US-10-109-551-2	Sequence 2, Appli	
6	1321	92.1	255	10	US-09-943-906-4	Sequence 4, Appli	
7	1316.5	91.7	256	9	US-10-109-551-4	Sequence 4, Appli	
8	1315.5	91.7	256	9	US-10-109-551-6	Sequence 6, Appli	
9	1315.5	91.7	256	9	US-10-109-551-8	Sequence 8, Appli	
10	1312.5	91.5	256	9	US-10-109-551-10	Sequence 10, Appl	
11	1299.5	90.6	254	9	US-10-106-574-5	Sequence 5, Appli	
12	1296.5	90.3	254	10	US-09-943-906-1	Sequence 1, Appli	
13	1293.5	90.1	254	9	US-10-106-574-6	Sequence 6, Appli	
14	1287.5	89.7	254	9	US-10-106-574-7	Sequence 7, Appli	
15	1281.5	89.3	254	9	US-10-106-574-8	Sequence 8, Appli	
16	1126	78.5	439	9	US-10-115-984-2	Sequence 2, Appli	
17	864	60.2	162	10	US-09-745-003-10	Sequence 10, Appl	
18	825	57.5	161	10	US-09-745-003-9	Sequence 9, Appli	
19	813	56.7	161	10	US-09-745-003-7	Sequence 7, Appli	

ALIGNMENTS

RESULT 1

```

US-09-904-987-3
; Sequence 3, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908Alacty
; TITLE OF INVENTION: Methods and Co
; TITLE OF INVENTION: Protein Assem
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 253
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI E
; DATABASE ENTRY DATE: 2001-04-17
; RELEVANT RESIDUES: (1)..(253)
US-09-904-987-3

```

	Query Match	100.0%	Score 1435;	DB 10;	Length 253;
	Best Local Similarity	100.0%;	Pred. No.	8.5e-117;	
	Matches 253;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MANLGCWMLVLFATWSDLGLCKKRPKGWGNTGGSRYPGQGSPGGNRRYPQGGGWGP	60		
Dd	1	MANLGCWMLVLFATWSDLGLCKKRPKGWGNTGGSRYPGQGSPGGNRRYPQGGGWGP	60		
Qy	61	HGGGWGPHGGGWQP HGGGWGQPHGGGWGGGTTHSQWNKP SKPKTNMKHMAAAAGA	120		
Dd	61	HGGGWGPHGGGWQP HGGGWGQPHGGGWGGGTTHSQWNKP SKPKTNMKHMAAAAGA	120		
Qy	121	VVGGLGYM LGSAMSRPI IHFGSDYEDRYYRENHNRRYPNQVVYRPMDEYSQNMFVHDCV	180		
Dd	121	VVGGLGYM LGSAMSRPI IHFGSDYEDRYYRENHNRRYPNQVVYRPMDEYSQNMFVHDCV	180		
Qy	181	NITTKQHVTVT TTTT KGENFTETDVKMMERVVEQMCI QYERESQAY YQRSSMWLFSSPPV	240		
Dd	181	NITTKQHVTVT TTTT KGENFTETDVKMMERVVEQMCI QYERESQAY YQRSSMWLFSSPPV	240		

QY 241 ILLISFLIFLIVG 253
 |||||
 Db 241 ILLISFLIFLIVG 253

RESULT 2

US-09-943-906-2
 : Sequence 2, Application US/09943906
 : Patent No. US20020150571A1
 : GENERAL INFORMATION:
 : APPLICANT: Prusiner, Stanley B.
 : Williamson, R. Anthony
 : Burton, Dennis R.
 : TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
 : NUMBER OF SEQUENCES: 86
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson P.C.
 : STREET: 2200 Sand Hill Road
 : CITY: Menlo Park
 : STATE: CA
 : COUNTRY: U.S.A.
 : ZIP: 94025
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/943,906
 : FILING DATE: 30-Aug-2001
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/550,374
 : FILING DATE: <Unknown>
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Bozicevic, Karl
 : REGISTRATION NUMBER: 28,807
 : REFERENCE/DOCKET NUMBER: 06510/059001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-854-5277
 : TELEFAX: 415-854-0875
 : TELEX: <Unknown>
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 253 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-943-906-2

Query Match 100.0%; Score 1435; DB 10; Length 253;
 Best Local Similarity 100.0%; Pred. No. 8.5e-117;
 Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
 |||||
 Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
 |||||
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
 |||||
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
 |||||
 QY 121 VVGLGGYMLGSAMSPRIIFHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 |||||
 Db 121 VVGLGGYMLGSAMSPRIIFHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 |||||
 QY 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 |||||
 Db 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 |||||

QY 241 ILLISFLIFLIVG 253
 |||||
 Db 241 ILLISFLIFLIVG 253

RESULT 3

US-09-919-172-57
 : Sequence 57, Application US/09919172
 : Patent No. US20020119463A1
 : GENERAL INFORMATION:
 : APPLICANT: Paris, Mary
 : APPLICANT: Turner, Christopher M.
 : TITLE OF INVENTION: PROSTATE CANCER MARKERS
 : FILE REFERENCE: PA-0036 US
 : CURRENT APPLICATION NUMBER: US/09/919,172
 : CURRENT FILING DATE: 2001-07-30
 : PRIOR APPLICATION NUMBER: 60/222,469
 : PRIOR FILING DATE: 2000-07-28
 : NUMBER OF SEQ ID NOS: 102
 : SOFTWARE: PERL Program
 : SEQ ID NO 57
 : LENGTH: 253
 : TYPE: PrT
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : OTHER INFORMATION: Incyte ID No. US20020119463A1 1256895CD1
 US-09-919-172-57

Query Match 99.7%; Score 1431; DB 10; Length 253;
 Best Local Similarity 99.6%; Pred. No. 1.9e-116;
 Matches 252; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
 |||||
 Db 1 MANLGCMVLVLFVATWSDGLCKKRPKPGGNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
 |||||
 QY 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
 |||||
 Db 61 HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAA 120
 |||||
 QY 121 VVGLGGYMLGSAMSPRIIFHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 |||||
 Db 121 VVGLGGYMLGSAMSPRIIFHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNFFVHDCV 180
 |||||
 QY 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 |||||
 Db 181 NITIKQHTVTTTKGENFTETDVKMERVVEQMCITQYERESQAYYQRGSSMVLFSPPV 240
 |||||
 QY 241 ILLISFLIFLIVG 253
 |||||
 Db 241 ILLISFLIFLIVG 253

RESULT 4

US-09-943-906-3
 : Sequence 3, Application US/09943906
 : Patent No. US20020150571A1
 : GENERAL INFORMATION:
 : APPLICANT: Prusiner, Stanley B.
 : Williamson, R. Anthony
 : Burton, Dennis R.
 : TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
 : NUMBER OF SEQUENCES: 86
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson P.C.
 : STREET: 2200 Sand Hill Road
 : CITY: Menlo Park
 : STATE: CA
 : COUNTRY: U.S.A.
 : ZIP: 94025
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette

QY 2 ANLGCWMLVLFVATWSDGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPGGGGWGQP 60
Db 4 SHIGSWILVLFVAMWSDVGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPGGGGWGQP 63
QY 61 HGGGWQPHGGGWGQPHGGGWGQPH-GGGWGGGGTHSQWNKPSKPKTNMKHMAGAAAAG 119
Db 64 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQ-GGTHSQWNKPSKPKTNMKHMAGAAAAG 122
QY 120 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDC 179
Db 123 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDC 182
QY 180 VNITKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPP 239
Db 183 VNITKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPP 242
QY 240 VILLISFLIFLIVG 253
Db 243 VILLISFLIFLIVG 256

RESULT 10
US-10-109-551-10
; Sequence 10, Application US/10109551
; Publication No. US20020194635A1
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; FILE REFERENCE: TAMK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Cervus elaphus
US-10-109-551-10

Query Match 91.5%; Score 1312.5; DB 9; Length 256;
Best Local Similarity 90.2%; Pred. No. 3.2e-106;
Matches 229; Conservative 19; Mismatches 3; Indels 3; Gaps 3;

QY 2 ANLGCWMLVLFVATWSDGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPGGGGWGQP 60
Db 4 SHIGSWILVLFVAMWSDVGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPGGGGWGQP 63
QY 61 HGGGWQPHGGGWGQPHGGGWGQPH-GGGWGGGGTHSQWNKPSKPKTNMKHMAGAAAAG 119
Db 64 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQ-GGTHSQWNKPSKPKTNMKHMAGAAAAG 122
QY 120 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDC 179
Db 123 AVVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDC 182
QY 180 VNITKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPP 239
Db 183 VNITKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFSPP 242
QY 240 VILLISFLIFLIVG 253
Db 243 VILLISFLIFLIVG 256

; APPLICANT: Harris, David A.
; APPLICANT: Stewart, Richard S.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-5

Query Match 90.6%; Score 1299.5; DB 9; Length 254;
Best Local Similarity 89.4%; Pred. No. 4.2e-105;
Matches 228; Conservative 15; Mismatches 9; Indels 3; Gaps 2;

QY 1 MANLGCWMLVLFVATWSDGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPGGGGWGQP 60
Db 1 MANLGYLLALFVMTWTDVGLCKRKP-KPGWNTGSRYPGQSPGGRNRYPPQ-GGTWGP 59
QY 61 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAG 120
Db 60 HGGGWQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAG 119
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDCV 180
Db 120 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMYRNQVYRPMDEYSNQNFFVHDCV 179
QY 181 NITIKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYY--ORGSSMVLFSPP 238
Db 180 NITIKQHTVTTTGTGENTFETDVKMMERVVEQMCITQYERESQAYY--ORGSSMVLFSPP 239
QY 239 PVILLISFLIFLIVG 253
Db 240 PVILLISFLIFLIVG 254

RESULT 12
US-09-943-906-1
; Sequence 1, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; Williamson, R. Anthony
; Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,906
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/550,374
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-943-906-1

Query Match 90.3%; Score 1296.5; DB 10; Length 254;
Best Local Similarity 89.0%; Pred. No. 7.6e-105;
Matches 227; Conservative 16; Mismatches 9; Indels 3; Gaps 2;

QY 1 MANLCWMLVAVATWSDLGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQGGGQGQP 60
DB 1 MANLGWLLALFVYMTDVGGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQ-GGTWQOP 59
QY 61 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 60 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 119
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 180
DB 120 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 179
QY 181 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYY--QRGSSMWLFSSP 238
DB 180 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYYDGRSSSTVLFSSP 239
QY 239 PVLLISFLIFLIVG 253
DB 240 PVLLISFLIFLIVG 254

RESULT 13

US-10-106-574-6
; Sequence 6, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion Dis
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-6

Query Match 90.1%; Score 1293.5; DB 9; Length 254;
Best Local Similarity 89.0%; Pred. No. 1.4e-104;
Matches 227; Conservative 15; Mismatches 10; Indels 3; Gaps 2;

QY 1 MANLCWMLVAVATWSDLGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQGGGQGQP 60
DB 1 MANLGWLLALFVYMTDVGGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQ-GGTWQOP 59
QY 61 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 60 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 119
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 180
DB 120 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 179

QY 181 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYY--QRGSSMWLFSSP 238
DB 180 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYYDGRSSSTVLFSSP 239
QY 239 PVLLISFLIFLIVG 253
DB 240 PVLLISFLIFLIVG 254

RESULT 14

US-10-106-574-7
; Sequence 7, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-7

Query Match 89.7%; Score 1287.5; DB 9; Length 254;
Best Local Similarity 88.2%; Pred. No. 4.5e-104;
Matches 225; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 1 MANLCWMLVAVATWSDLGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQGGGQGQP 60
DB 1 MANLGWLLALFVYMTDVGGLCKRKPFGGWTGSGRYPGQSPGNGNRYPPQ-GGTWQOP 59
QY 61 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 120
DB 60 HGGGWQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPTNNKHMAGAAAAGA 119
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 180
DB 120 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNQNNEFVHDCV 179
QY 181 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYY--QRGSSMWLFSSP 238
DB 180 NITIKQHTVTTTNGENFTETDVKKMERVVEQMCITQYERESQAYYDGRSSSTVLFSSP 239
QY 239 PVLLISFLIFLIVG 253
DB 240 PVLLISFLIFLIVG 254

RESULT 15

US-10-106-574-8
; Sequence 8, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-8

Query Match 89.3%; Score 1281.5; DB 9; Length 254;

Best Local Similarity 87.8%; Pred. No. 1.5e-103;
Matches 224; Conservative 15; Mismatches 13;

1	MANLGCMWLVL	FVATWSDGLCKKRP	KPGGWNTGSRYP	PGGSPGGRNRY	PPQGGGQWGP	60
Qy						
1	MANLGWLRLA	LEFTWTDVGLCKKRP	KPGGMWTTGSRYP	PGGSPGGRNRY	PPQ--GTTWQP	59
Db						
61	HGGWGQGP	HGGGQWGP	HGGGQWGP	HGGGQWGP	HGGGQWGP	120
Qy						
60	HGGWGQGP	HGGGQWGP	HGGGQWGP	HGGGQWGP	HGGGQWGP	119
Db						
121	VVGLGGYMLGS	AMSRIPIHFGSDY	EDRYRENMRHY	PQVYRPDMEY	SQNNNFVHDCV	180
Qy						
120	VVGLGGYMLGS	AVSRPMIHFGNDW	EDRYRENMYR	PQVYRPVQYS	QNQNFVHDCV	179
Db						
181	NITTKQHTV	TTTTKGFNTETD	VKMVERVEQMC	ITQYERESQAY	--ORGSSWLFSSP	238
Qy						
180	NITTKQHTV	TTTTKGFNTETD	VKMVERVEQMC	ITQYERESQAY	YDGRSSSTVLFPSP	239
Db						
239	PVILLISFLI	FLIV	253			
Qy						
240	PVILLISFLI	FLIV	254			
Db						

Search completed: January 28, 2003, 14:40:07
Job time : 4.31713 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 : Search time 5.25212 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-3

Perfect score: 1435

Sequence: 1 MANLCGWLVLFPVATWSDLG.....LFSSPPVILLISFLIFLIVG 253

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1435	100.0	253	1	US-08-242-188-2
2	1435	100.0	253	1	US-08-509-261A-2
3	1435	100.0	253	1	US-08-660-626-8
4	1435	100.0	253	1	US-08-692-892-2
5	1435	100.0	253	2	US-08-713-939A-2
6	1435	100.0	253	2	US-08-868-162A-22
7	1435	100.0	253	4	US-09-031-168-8
8	1435	100.0	253	4	US-09-128-450-20
9	1435	100.0	253	4	US-09-036-579-2
10	1435	100.0	253	4	US-09-823-494-20
11	1435	100.0	253	4	US-09-550-374-2
12	1331.5	92.8	264	4	US-09-128-450-21
13	1331.5	92.8	264	4	US-09-823-494-21
14	1329	92.6	263	1	US-08-242-188-3
15	1329	92.6	263	1	US-08-509-261A-3
16	1329	92.6	263	1	US-08-660-626-9
17	1329	92.6	263	1	US-08-692-892-3
18	1329	92.6	263	2	US-08-713-939A-3
19	1329	92.6	263	2	US-08-868-162A-23
20	1329	92.6	263	4	US-09-031-168-9
21	1329	92.6	263	4	US-09-550-374-3
22	1329	92.6	263	4	US-09-823-494-21
23	1321	92.1	255	1	US-08-242-188-4
24	1321	92.1	255	1	US-08-509-261A-4
25	1321	92.1	255	1	US-08-660-626-10
26	1321	92.1	255	1	US-08-692-892-4
27	1321	92.1	255	2	US-08-713-939A-4

28	1321	92.1	255	2	US-08-868-162A-24
29	1321	92.1	255	4	US-09-031-168-10
30	1321	92.1	255	4	US-09-036-579-4
31	1321	92.1	255	4	US-09-550-374-4
32	1314.5	91.6	256	4	US-09-128-450-22
33	1314.5	91.6	256	4	US-09-823-494-22
34	1308.5	91.2	254	4	US-09-128-450-26
35	1308.5	91.2	254	4	US-09-823-494-26
36	1296.5	90.3	254	1	US-08-242-188-1
37	1296.5	90.3	254	1	US-08-509-261A-1
38	1296.5	90.3	254	1	US-08-660-626-7
39	1296.5	90.3	254	1	US-08-692-892-1
40	1296.5	90.3	254	2	US-08-713-939A-1
41	1296.5	90.3	254	2	US-08-868-162A-21
42	1296.5	90.3	254	4	US-09-031-168-7
43	1296.5	90.3	254	4	US-09-128-450-19
44	1296.5	90.3	254	4	US-09-036-579-1
45	1296.5	90.3	254	4	US-09-823-494-19

ALIGNMENTS

RESULT 1

US-08-242-188-2
; Sequence 2, Application US/08242188
; Patent No. 5565186
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: METHOD OF DETECTING PRIONS IN A SAMPLE
; TITLE OF INVENTION: AND TRANSGENIC ANIMAL USED FOR SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bosicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,188
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bosicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/014001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-5277
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrp
US-08-242-188-2

Query Match 100.0%; Score 1435; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLCGWLVLFPVATWSDLGLCKKRPKPGWNTGSGRYPGQSPGCGNRYPPQGGGWCQP 60

Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLFSPPV 240
Db 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 4

US-08-692-892-2
; Sequence 2, Application US/08692892
; Patent No. 5792901
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Scott, Michael R.
; APPLICANT: Telling, Glenn
; TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND
; TITLE OF INVENTION: PRION PREPARATION AND TRANSGENIC ANIMAL USED FOR SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Karl Bozicevic
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/692,892
; FILING DATE: 30-JULY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-08-692-892-2

Query Match 100.0%; Score 1435; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGGWNTPGSGSPGNRYPPQGGGGWGQP 60
Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGGWNTPGSGSPGNRYPPQGGGGWGQP 60
QY 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLFSPPV 240
Db 181 NITIKOHTVTTTNGENFTETDVKMERVVEQMCITQYERESQAYQRGSSMVLFSPPV 240
QY 241 ILLISFLIFLIVG 253
Db 241 ILLISFLIFLIVG 253

RESULT 5

US-08-713-939A-2
; Sequence 2, Application US/08713939A
; Patent No. 5846533
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,939A
; FILING DATE: 13-SEP-1996
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-713-939A-2

Query Match 100.0%; Score 1435; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCWMLVLFVATWSDGLCKKRPKPGGWNTPGSGSPGNRYPPQGGGGWGQP 60
Db 1 MANLGCWMLVLFVATWSDGLCKKRPKPGGWNTPGSGSPGNRYPPQGGGGWGQP 60
QY 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
Db 61 HGGGWCQPHGGGQPHGGGQGGTSHQWKNKPKTKNMHAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHRYPNQVYRPMDEYSNONNFVHDCV 180

Db	121	VVGGLGMYLGSAMSRPIIHFGSDYEDRYRRENNHRRPNQVYRPMDEYTSNQNNFVHDCV	180
Qy	181	NITIKQHTVTTTKGNETETDKMMERVVEQNCITQYERESQAYYQRGSSMVLFSPPV	240
Db	181	NITIKQHTVTTTKGNETETDKMMERVVEQNCITQYERESQAYYQRGSSMVLFSPPV	240
Qy	241	ILLISFLIFLIVG	253
Db	241	ILLISFLIFLIVG	253

```

RESULT 6
US-08-868-162A-22
; Sequence 22, Application US/08868162A
; Patent No. 5962569
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley
; APPLICANT: Cohen, Fred
; APPLICANT: James, Thomas
; APPLICANT: Kaneko, Kiyotoshi
; TITLE OF INVENTION: Prion Protein Modulator Factor
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,162A
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 6510-083001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
; US-08-868-162A-22

```

```
Query Match      100.0%; Score 1435; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1	MANLGCWMLVLFVATWSDSLGLCKKPKPGCWNTGGSRYPGQSGPGNRYPPGGGGWGQ	60
Qy		
1	MANLGCWMLVLFVATWSDSLGLCKKPKPGCWNTGGSRYPGQSGPGNRYPPGGGGWGQ	60
Db		
1	MANLGCWMLVLFVATWSDSLGLCKKPKPGCWNTGGSRYPGQSGPGNRYPPGGGGWGQ	60
Db		
61	GGGGWQPHGGGWGQPHGGGGWQPHGGGGWQGGGTHSQNNKPSKPTNNKMHAGAAAAA	120
Qy		
61	HGGGWQPHGGGWGQPHGGGGWQPHGGGGWQGGGTHSQNNKPSKPTNNKMHAGAAAAA	120
Db		

[illegible]

```

RESULT 7
US-09-031-168-8
; Sequence 8, Application US/09031168
; Patent No. 6150583
; GENERAL INFORMATION:
; APPLICANT: Stanley B. Prusiner
; APPLICANT: Glenn E. Telling
; APPLICANT: Fred E. Cohen
; APPLICANT: Michael R. Scott
; TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING
; TITLE OF INVENTION: EPIOTOPE-TAGGED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSED: Fish & Richardson
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/660,626
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Valeta Gregg
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 07533/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN PRION PROTEIN, HuPrP
US-09-031-168-8

```

```
Query Match      100.0%; Score 1435; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1	Qy	MANLGCWMLVLVFVATWSDLGCLCKRRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQP	60
1	Db	MANLGCWMLVLVFVATWSDLGCLCKRRPKPGWNTGSGRYPGQSPGNGNRYPPQGGGGWGQP	60
61	Qy	HGGGGQPHGGGWAQPHGGGGWGQPHGGGGGQGGGTHSQWNPKSPKPTNKKIMACAAAAGA	120
61	Db	HGGGGQPHGGGGWGQPHGGGGWGQPHGGGGGCGGGTHSQWNPKSPKPTNKKIMACAAAAGA	120

QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
|||||
Db 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 8

US-09-128-450-20
; Sequence 20, Application US/09128450
; Patent No. 6211149
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Sulette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/128,450
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-128-450-20

Query Match 100.0%; Score 1435; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
|||||
Db 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNMKHMAGAAAAGA 120
|||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNMKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
|||||
Db 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 9

US-09-036-579-2
; Sequence 2, Application US/09036579
; Patent No. 6230954
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; APPLICANT: Williamson, R. Anthony
; APPLICANT: Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.

; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,579
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,939
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-579-2

Query Match 100.0%; Score 1435; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2e-133;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
|||||
Db 1 MANLGCMVLVLFVATWSDLGLCKKRPKPGGWNTGSGRYPGQSGPGNRYPPQGGGGWGQP 60
QY 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNMKHMAGAAAAGA 120
|||||
Db 61 HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGHGTHSQWNKPSKPTNMKHMAGAAAAGA 120
QY 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
|||||
Db 121 VVGLGGYMLGSAMSRPIIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQNNFVHDCV 180
QY 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
|||||
Db 181 NITIKOHTVTTTNGENFTETDVKMMRVVEQMCITQYERESQAYYQRGSSMWLFSSPPV 240
QY 241 ILLISFLIFLIVG 253
|||||
Db 241 ILLISFLIFLIVG 253

RESULT 10

US-09-823-494-20
; Sequence 20, Application US/09823494
; Patent No. 6355610
; GENERAL INFORMATION:
; APPLICANT: Chesebro, Bruce W
; APPLICANT: Caughey, Byron W
; APPLICANT: Chabry, Joelle
; APPLICANT: Priola, Sulette
; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion
; FILE REFERENCE: 50121
; CURRENT APPLICATION NUMBER: US/09/823,494
; CURRENT FILING DATE: 2001-03-30


```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/509/261A
; FILING DATE: 31-JUL-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 6510-030001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-509-261A-3

Query Match          92.6%; Score 1329; DB 1; Length 263;
Best Local Similarity 87.7%; Pred. No. 5.5e-123;
Matches 228; Conservative 19; Mismatches 5; Indels 8; Gaps 1;

QY  2  ANLCGWMLVLFVATWSDGLCKKRPKPGGNTGSGRYPGGSPGGNRYPPQGGGQGP 61
Db   4  SHIGSWILVLFVAMWSDVGLCKKRPKPGGNTGSGRYPGGSPGGNRYPPQGGGQGP 63
QY  62  GGGWGQPHGGGQGP 113
Db   64  GGGWGQPHGGGQGP 123
QY  114  GAAAGAVGGLGGLGMLGMSAMRPIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQ 173
Db   124  GAAAGAVGGLGGLGMLGMSAMRPLIHFGSDYEDRYRENHMRYPNQVYRPMDEYSNQ 183
QY  174  NFVHDCVNITIKOHTVTTTGTGENFTETDVKMERVVEQMCITQYERESQAYYQGSVMV 233
Db   184  NFVHDCVNITVKEHTVTTTGTGENFTETDVKMERVVEQMCITQYERESQAYYQGSVMV 243
QY  234  LFSPPVILLISFLIFLIV 253
Db   244  LFSPPVILLISFLIFLIV 263
```

Search completed: January 28, 2003, 14:21:21
Job time : 6.25212 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 3.67113 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKGEGVVAEAE.....DNEAYEMPSEEGYQDYEPPEA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	702	100.0	140	2 A49669	synuclein alpha -
2	485	69.1	149	2 B43959	synuclein, form 2
3	409	58.3	134	2 S44430	synuclein - human
4	408	58.1	134	2 S39045	phosphoneuroprotei
5	407	58.0	137	2 I56498	synuclein-like pro
6	273.5	39.0	123	2 A57431	synuclein - Pacifi
7	261	37.2	113	2 A60887	14K brain-specific
8	245.5	35.0	124	2 A60218	alpha-synuclein, N
9	129	18.4	28	2 S56746	protein 135, daue
10	123	17.5	150	2 T37479	hypothetical prote
11	122	17.4	798	2 T21369	seed biotin-contai
12	105	15.0	643	1 T07064	embryonic abundan
13	100	14.2	317	2 S36116	embryonic abundant
14	94	13.4	224	2 S16259	probable peptidogl
15	90.5	12.9	1612	2 AB1347	pectinesterase hom
16	89	12.7	474	2 T01347	methyl-accepting c
17	89	12.7	739	2 H75001	surface-array prot
18	89	12.7	1109	2 A36143	membrane protease
19	88.5	12.6	322	2 A37030	CAMP-dependent pro
20	88	12.5	464	2 A32461	embryonic abundant
21	87.5	12.5	473	2 S61428	embryonic protein
22	87.5	12.5	474	2 S39475	phosphoribosylform
23	87.5	12.5	742	1 S78526	phosphoribosylamin
24	87	12.4	424	2 B70365	phage sheath prote
25	87	12.4	469	2 E81121	pyruvate, water di
26	87	12.4	817	2 J04176	tail protein (bact
27	87	12.4	1093	2 AC1753	probable chemotaxi
28	86	12.3	739	2 F71161	major merozoite su
29	86	12.3	1785	2 A45546	

30 85.5 12.2 448 2 H84782 late embryogenesis
31 85 12.1 288 2 S38219 ABA-inducible prot
32 85 12.1 791 2 A53206 6-phosphofructokin
33 85 12.1 2038 2 A43742 female sterile hom
34 84.5 12.0 350 2 S61426 late embryogenesis
35 84.5 12.0 1467 2 A75564 conserved hypothet
36 84 12.0 216 2 S49047 vmp26 protein - Bo
37 84 12.0 1619 2 D97132 uncharacterized ph
38 83.5 11.9 384 2 G86287 hypothetrical prote
39 83.5 11.9 479 2 T47561 late embryogenesis
40 83.5 11.9 480 2 C96744 hypothetrical prote
41 83.5 11.9 894 2 E82221 DNA gyrase, chain
42 83.5 11.9 1787 2 AG1360 probable tape-meas
43 82.5 11.8 280 2 S04130 embryonic abundant
44 82.5 11.8 448 2 JC6171 late embryogenesis
45 82.5 11.8 458 2 S33520 Lea protein - soyb

ALIGNMENTS

RESULT 1

A49669

synuclein alpha - human

N:Alternate names: non-A4/beta component, Alzheimer disease amyloid plaque (NAC)
C:Species: Homo sapiens (man)

C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 05-Nov-1999

C:Accession: A49669

R:Ueda, K.; Fukushima, H.; Masiliah, E.; Xia, Y.; Iwai, A.; Yoshimoto, M.; Otero, D.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 11282-11286, 1993

A:Title: Molecular cloning of cDNA encoding an unrecognized component of amyloid in A
A:Reference number: A49669; MUID:94068588; PMID:8248242

A:Accession: A49669

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <RES>

A:Cross-references: GB:I08850; NID:9437364; PIDN:AAA16117.1; PID:9437365

A:Note: parts of this sequence were determined by protein sequencing

C:Genetics:

A:Gene: GDB:SNCA

A:Cross-references: GDB:439047; OMIM:163890

A:Map position: 4q21-4q21

C:Keywords: amyloid

Query Match 100.0%; Score 702; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.3e-49;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDVFMKGLSKAKGEGVVAEAEKTGCGVAEAKTKGEGVLYVGSKTKEGVVHGVAETVAETK 60

Db 1 MDVFMKGLSKAKGEGVVAEAEKTGCGVAEAKTKGEGVLYVGSKTKEGVVHGVAETVAETK 60

Qy 61 EQVTNVGGAVVTGVTAVAAKTVEGAGSAAATGFKKDLQCKNEGAPQEGILEDMPVDP 120

Db 61 EQVTNVGGAVVTGVTAVAAKTVEGAGSAAATGFKKDLQCKNEGAPQEGILEDMPVDP 120

Qy 121 DNEAYEMPSEEGYQDYEPPEA 140

Db 121 DNEAYEMPSEEGYQDYEPPEA 140

RESULT 2

B43959

synuclein, form 2 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Dec-2000

C:Accession: B43959; A43959; C43959; B60887

R:Maroteaux, L.; Scheller, R.H.

Brain Res. Mol. Brain Res. 11, 335-343, 1991

A:Title: The rat brain synucleins; family of proteins transiently associated with neu

A:Reference number: A43959; MUID:92099853; PMID:1661825

A:Accession: B43959

A:Molecule type: mRNA

synuclein-like protein - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 05-Nov-1999
 C:Accession: A57431

R:AKopian, A.N.; Wood, J.N.

J. Biol. Chem. 270, 21264-21270, 1995

A:Title: Peripheral nervous system-specific genes identified by subtractive cDNA cloning

A:Reference number: A57431

A:Accession: A57431

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-123 <AKO>

A:Cross-references: GB:X86789; NID:g971579; PIDN:CAA60485.1; PID:g971580

Query Match 39.0%; Score 273.5; DB 2; Length 123;
 Best Local Similarity 55.8%; Pred. No. 3.1e-15;
 Matches 63; Conservative 12; Mismatches 35; Indels 3; Gaps 1;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKEGVLYVSGTKEGVYVGHVATVAEKTK 60

Db. 1 MDVFKKGFSTAREGVVGAETKQGVAAAEKTKEGVYVGHVATVAEKTK 60

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

QY 61 EQVTNVGGAVVGTAVAAKTVGAGSIAAATGFVKKDQL---GKNEEGAPQE 110

Db 61 EQANVSEAVSVNVATVETVEAENIVVTGVVRKEDLEPPAQDQEAKEQE 113

A:Residues: 1-124 <NAK>

Query Match

Best Local Similarity 35.0%; Score 245.5; DB 2; Length 124;

Matches 69; Conservative 15; Mismatches 21; Indels 35; Gaps 8;

QY 1 MDVEMK-GLSKA--KEGVVAAAEKTKQGVAAAGTKEGVLYVSGTKEGVYVGHVATVAE 57

Db 1 MDVFMKQGVTEAAEKEGVVAAAEK---GLSMA-----KEGVLYVGS--KEGVVQGVASVAE 51

QY 58 KTKEQVTNVGGAVVGT---VTAAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 52 KTKEQASHLGGAVFSSAGNEQAASHLGGAVFSSAGNIAAATGLVKK----- 96

QY 113 LEDMPVD--PDNEAYEMPSE 130

Db 97 -EEFPTDLKPEEVAQEAEE 115

RESULT 9

S56746

alpha-synuclein, NAC - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000

C:Accession: S56746

R:Jensen, P.H.; Sorensen, E.S.; Petersen, T.E.; Gliemann, J.; Rasmussen, L.K.

Biochem. J. 310, 91-94, 1995

A:Title: Residues in the synuclein consensus motif of the alpha-synuclein fragment, N

A:Reference number: S56746

A:Accession: S56746

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-28 <JEN>

Query Match 18.4%; Score 129; DB 2; Length 28;

Best Local Similarity 100.0%; Pred. No. 0.00025;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EQVTNVGGAVVGTAVAAKTVGAGS 87

Db 2 EQVTNVGGAVVGTAVAAKTVGAGS 28

RESULT 10

T37479

Protein 135, dauer-upregulated - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T37479

R:Cherkasova, V.A.; Egilmez, N.K.; Shmookler Reis, R.J.

submitted to the EMBL Data Library, March 1996

A:Description: Diverse Caenorhabditis elegans genes that are upregulated in dauer lar

A:Reference number: Z21715

A:Accession: T37479

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-150 <CHE>

A:Cross-references: EMBL:U52071; PIDN:AAD11528.1

A:Experimental source: strain Bristol-N2

C:Genetics:

A:Gene: dur135

A:Map position: IV

Query Match

Best Local Similarity 17.5%; Score 123; DB 2; Length 150;

Matches 36; Conservative 12; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVVAAAEKTKQGVAAAEKTKEGVLYVSGTKEGVYVGHVATVAEKT---EQVTNVGGA 69

Db 10 DGAVAAKDAVEGATTAAGTKTEGMRFG----ENVAHGAGKAVADAKAAGDTVVEGKA 65

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

QY 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 70 VV----TGVTAVAAKTVGAGSIAAATGFVKKDQLGKNEEGAPQEGI 112

Db 66 AAEKTESGIAFVGOKIAEGAGAVAGG-----AKAAGGAVVDGI 103

RESULT 11

T21369 hypothetical protein F25H8.5a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21369

R:Gajadaty, S.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z19413

A:Accession: T21369

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-796 <WIL>

A:Cross-references: EMBL:Z69360; PIDN:CAA93285.1; GSPDB:GN00022; CESP:F25H8.5a

A:Experimental source: clone F25H8

C:Genetics:

A:Gene: CESP:F25H8.5a

A:Map position: 4

A:Introns: 23/3; 43/1; 303/1; 321/1; 343/1; 380/1; 387/1; 405/1; 427/1; 449/1; 478/1; 49

Query Match 17.4%; Score 122; DB 2; Length 798;

Best Local Similarity 32.7%; Pred. No. 0.035;

Matches 35; Conservative 13; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVVAAAEKTKQGVAAEAKTKRGVLYVSGTKGVVGHVATVAETK---EQVTNVGGA 69

Db 610 DGAVAAKDAADAAEATAGNKTVECMRFV-----ENVAHGAGKVADGAKAAGDVTVEGGA 665

QY 70 VV-----TGVTAAVAKTVEGAGSATAATGFVKDQLGKNEEAPQEGT 112

Db 666 AAEKTESGIAFVGOKIAEGAGAVAGG-----AKAAGGAVVDGV 703

RESULT 12

T07064

seed biotin-containing protein LEA [validated] - soybean

C:Species: Glycine max (soybean)

C:Date: 01-Sep-2000 #sequence_revision 01-Sep-2000 #text_change 01-Feb-2002

C:Accession: T07064

R:Hsing, Y.C.; Tsou, C.H.; Hsu, T.F.; Chen, Z.Y.; Hsieh, K.L.; Hsieh, J.S.; Chow, T.Y.

Plant Mol. Biol. 38, 461-490, 1998

A:Title: Tissue- and stage-specific expression of a soybean (Glycine max L.) seed-mature

A:Reference number: Z15895; MUID:98418627; PMID:9747855

A:Accession: T07064

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-643 <HSI>

A:Cross-references: EMBL:U59626; NID:gl389896; PIDN:AAC61783.1; PID:gl389897

A:Experimental source: strain Shi-Shi; cotyledon

C:Superfamily: pea seed biotin-containing protein

C:Keywords: biotin binding; seed

F:125/Binding site: biotin (Lys) (covalent) #status predicted

Query Match 15.0%; Score 105; DB 1; Length 643;

Best Local Similarity 28.2%; Pred. No. 0.65;

Matches 48; Conservative 10; Mismatches 68; Indels 44; Gaps 7;

QY 2 DVFMKGLSKAKEGVVAAAEKTKQGVAAEAKTKRGVLYVSGTKGVVGHVATVAETK--- 51

Db 289 DTISAAKTAASEKTAPVAEAKADYTLQAEKAKSAGGTTTASYGKAVQAKDVAVESGKS 348

QY 52 ----VATVAEKTKEQVTNVGGA-----VVTGVTAAVAKTVEGAG----- 86

Db 349 AAGYAAKVAADLRDKATAVGAAAHFAEKTVEGTAKAAH-VVEGAGYAGHKAELASM 407

QY 87 -----STAAATGFVKKDLQKNEEAPQEGILEMDPVDPNAYEMPSE 130

Db 40b SAGAVKGLAASAGETAKEYTAKKKEAQRE--LEAKKPSQPQAEERPSE 455

RESULT 13

S33616

embryonic abundant protein, group 3 - wheat (fragment)

N:Alternate names: group 3 late embryogenesis abundant protein

C:Species: Triticum aestivum (common wheat)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Oct-1999

C:Accession: S33616; S27767

R:Curry, J.; Walker-Simmons, M.K.

Plant Mol. Biol. 21, 907-912, 1993

A:Title: Unusual sequence of group 3 LEA (II) mRNA inducible by dehydration stress in

A:Reference number: S33616; MUID:93222487; PMID:8467082

A:Accession: S33616

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-317 <WAL>

A:Cross-references: EMBL:M72395; NID:gl70691; PIDN:AAA34267.1; PID:gl70692

Query Match 14.2%; Score 100; DB 2; Length 317;

Best Local Similarity 29.6%; Pred. No. 0.75;

Matches 53; Conservative 10; Mismatches 62; Indels 54; Gaps 8;

QY 6 KGLSKAKEGVVAAAEKTKQGVAAE-----AAGKTREGVLYVSGTKGVVGHVATVAE 57

Db 9 EKEGASEMADAAGTKDKAAAEKTRMADAAAGTKRETAKDAAVEKTRRE-MADTAATKAA 67

QY 58 KTK-----BQVT-----NVGGAVVT-----GVTAQAQKTVEGA--- 85

Db 68 ETKDAAAEKASGAGEMVTEKARSADAAADKASGAETVTEKAKGAKDAALDTAEGAKEY 127

QY 86 -----GSTAANTGVKKDLQKNEEAPQEGILEMDPVDPNAYEMPSEEGVDYEP 138

Db 128 MVDKKEDARRALAGSV-KDSKGETNESACQGG-----QDVRRRAAEKAEARQTHOP 179

RESULT 14

S16259

embryonic abundant protein, group 3 - common wheat x Sanduri wheat

N:Alternate names: late embryogenesis abundant protein

C:Species: Triticum aestivum x Triticum timopheevi (common wheat x Sanduri wheat)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Sep-2000

C:Accession: S16259

R:Curry, J.; Morris, C.F.; Walker-Simmons, M.K.

Plant Mol. Biol. 16, 1073-1076, 1991

A:Title: Sequence analysis of a cDNA encoding a Group 3 LEA mRNA inducible by ABA or

A:Reference number: S16259; MUID:91322498; PMID:1830822

A:Accession: S16259

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-224 <CUR>

A:Cross-references: EMBL:X55882; NID:921636; PIDN:CAA40204.1; PID:g21637

C:Superfamily: barley hypothetical protein (clone ES2A)

Query Match 13.4%; Score 94; DB 2; Length 224;

Best Local Similarity 30.2%; Pred. No. 1.6;

Matches 35; Conservative 12; Mismatches 55; Indels 14; Gaps 4;

QY 6 KGLSKAKEGVVAAAEKTKQGVAAEAKTKRGVLYVSGTKGVVGHVATVAETK 60

Db 16 KARNEEKTGVGMGATKDKAGQTTEATKQAGETTEATKQAAETTEAAKQKASETAEATK 75

QY 61 EQVTNVGAVVTGVTAAVAKTVEGAGSI--AAATGFVKKQDQ----LGKNEGAPOE 110

Db 76 QKAAEAKDKTAQAAAEKTKYETAQSAKERAQGS---KQDTASTLGKTEAAKQK 128

RESULT 15

AB1347

probable peptidoglycan bound protein (LPXTG motif) lmo2178 [Imported] - Listeria mono

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1347

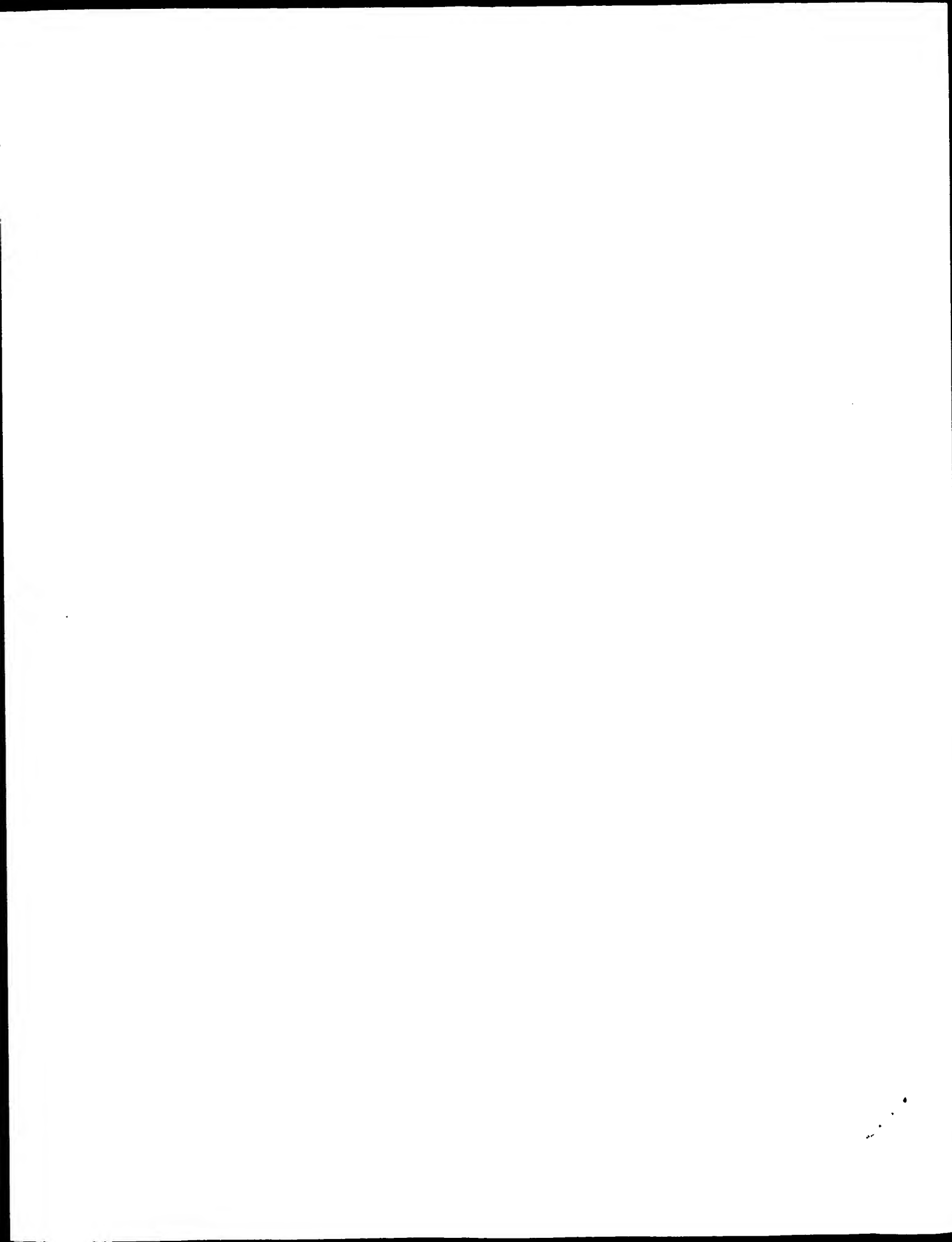
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.: Jones, L.M.: Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wenland,
A:Title: Comparative genomics of *Listeria* species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1347
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1612 <GLA>
A:Cross-references: GB:NC_003210; PID:CAD00256.1; PID:gl6411648; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2178

Query Match 12.9%; Score 90.5; DB 2; Length 1612;
Best Local Similarity 28.1%; Pred. No. 26;
Matches 45; Conservative 17; Mismatches 65; Indels 33; Gaps 7;
QY 3 VFMKGLSKAKGVAA-----AEKTKQG---VAEAAAGK-----TKEGVL 38
Db 1110 VLLKEDSATKDATAGAEFELOQADGKTVADNLVSNADGKIEVTDLPAGDYQFVETKAPTG 1169
QY 39 YV--GSKTKEGVYHGVATVAEKTKEQVTVNGGAVVTGVTVAQKTVGAG-SIAAATGEV 95
Db 1170 YVLDGAPTKEVTFEFOEAAVITKENTAKTGSVLTKEDSVSKATISGAEEFELQATGK 1229
QY 96 KKDQKNEEGAPQEGILEDMPVDP-DNEAYEMPSEEGYO 134
Db 1230 VKDNLTTNADGK-----LEVTDLPAGDYKFVETKAPTGYE 1264

Search completed: January 28, 2003, 14:20:02
Job time : 5.67113 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 2.14149 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKGYVAAAE.....DNEAYEMPSEGYQDYEPFA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	702	100.0	1 SYUA_HUMAN	P37840 homo sapien
2	670	95.4	1 SYUA_RAT	P37377 rattus norv
3	669	95.3	1 SYUA_MOUSE	O55042 mus musculu
4	595.5	84.8	1 SYUA_POEGU	Q91448 poephila gu
5	409	58.3	1 SYUB_HUMAN	Q16143 homo sapien
6	408	58.1	1 SYUB_BOVIN	P33567 bos taurus
7	407	58.0	1 SYUB_RAT	Q63754 rattus norv
8	337	48.0	1 SYU_TORCA	P37379 torpedo cal
9	313	44.6	1 SYUG_HUMAN	O76070 homo sapien
10	312.5	44.5	1 SYUG_MOUSE	Q920f7 mus musculu
11	307.5	43.8	1 SYUG_HUMAN	Q9n250 homo sapien
12	273.5	39.0	1 SYUG_RAT	Q63544 rattus norv
13	94	13.4	1 LEA3_WHEAT	Q03968 triticum ae
14	88	12.5	1 AK15_RAT	Q03968 triticum ae
15	87.5	12.5	1 PURL_BACSU	P24587 rattus norv
16	87	12.4	1 PURL_AQUAE	P12042 bacillus su
17	87	12.4	1 PPSA_PYREF	O66949 aquifex ao
18	86.5	12.3	1 P60_LISIV	P42850 pyrococcus
19	85	12.1	1 K6PP_RABIT	Q01837 listeria iv
20	85	12.1	1 FSH_DROME	P47859 ocytolaqus
21	84.5	12.0	1 P60_LISWE	P37099 drosophila
22	82.5	11.8	1 LE76_BRANA	Q1839 listeria we
23	82.5	11.8	1 SPAA_STRDO	P13934 brassica na
24	82	11.7	1 EAEB_ECO27	P1979 streptococ
25	82	11.7	1 HPR2_HALVA	Q05129 escherichia
26	82	11.7	1 POLG_FEMVM	P42258 haloarcula
27	81.5	11.6	1 LE11_THETN	O56075 p genome po
28	81	11.5	1 AMPH_CHICK	Q84k3 thermocoe
29	80.5	11.5	1 NT6A_HUMAN	P50478 gallus gall
30	79.5	11.3	1 YD39_HAEIN	P34132 homo sapien
31	79.5	11.3	1 PPSA_PYRAB	P71378 haemophilus
32	79	11.3	1 DRPF_CRAPL	Q92h7 pyrococcus
33	79	11.3	1 SYE_PYRHO	P3283 craterostig
				O59314 pyrococcus

RESULT 1

ID	SYUA_HUMAN	STANDARD;	PRT;	140 AA.
AC	P37840; Q13701;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component of amyloid precursor) (NACP).			
DE	SNCA OR NACP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND SEQUENCE OF 61-95.			
RC	TISSUE=Brain;			
RX	MEDLINE=94068588; PubMed=8248242;			
RA	Ueda K., Fukushima H., Masliak E., Xia Y., Iwai A., Yoshimoto M., Otero D.A., Kondo J., Ihara Y., Saitoh T.;			
RA	"Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 90:11282-11286(1993).			
RL	[2]			
RN	SEQUENCE FROM N.A. (ISOFORMS 1; 2-4 AND 2-5).			
RP	MEDLINE=95324916; PubMed=7601450;			
RX	Campion D., Martin C., Heilig R., Charbonnier F., Moreau V., Flaman J.-M., Petit J.-L., Hannequin D., Brice A., Frebourg T.;			
RA	"The NACP/synuclein gene: chromosomal assignment and screening for alterations in Alzheimer disease.";			
RT	Genomics 26:254-257(1995).			
RL	[3]			
RN	SEQUENCE FROM N.A. (ISOFORM 2-4).			
RP	TISSUE=Brain;			
RX	MEDLINE=95100969; PubMed=7802671;			
RA	Ueda K., Saitoh T., Mori H.;			
RT	"Tissue-dependent alternative splicing of mRNA for NACP, the precursor of non-A beta component of Alzheimer's disease amyloid.";			
RL	Biochem. Biophys. Res. Commun. 205:1366-1372(1994).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Xia Y., Silva R.D., Chen X.H., Saitoh T.;			
RL	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	PHOSPHORYLATION BY CASEIN KINASE.			
RX	MEDLINE=20085057; PubMed=10617630;			
RA	Okochi M., Walter J., Koyama A., Nakajo S., Baba M., Iwatsubo T., Meljer L., Kahle P.J., Haass C.;			
RT	"Constitutive phosphorylation of the Parkinson's disease associated alpha-synuclein.";			
RL	J. Biol. Chem. 275:390-397(2000).			
RN	[6]			
RP	PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR KINASE.			
RX	MEDLINE=20409007; PubMed=10852916;			
RA	Pronin A.N., Morris A.J., Surguchov A., Benovic J.L.;			
RT	"Synucleins are a novel class of substrates for G protein-coupled			

34	79	11.3	625	1	ILVB_MYCLE	O33112 mycobacteri
35	79	11.3	964	1	LON2_MAIZE	P3648 zea mays (m
36	78.5	11.2	739	1	PURL_LISIN	Q92an9 listeria in
37	78.5	11.2	739	1	PURL_LISMO	Q9y6c1 listeria in
38	77.5	11.0	213	1	LEA1_HORVU	P14928 hordeum vul
39	77.5	11.0	216	1	MBC1_METWA	P58979 methanosarc
40	77.5	11.0	221	1	LEA3_MAIZE	O42376 zea mays (m
41	77.5	11.0	280	1	MACS_CHICK	P16527 gallus gall
42	77.5	11.0	503	1	NRF1_HUMAN	Q16656 homo sapien
43	77.5	11.0	503	1	NRF1_MOUSE	Q9wu00 mus musculu
44	77.5	11.0	729	1	PURL_STAM	Q99v28 staphylococ
45	77	11.0	306	1	YOPD_YEREN	P37132 yersinia en

ALIGNMENTS

receptor kinases.";
 RL J. Biol. Chem. 275:26515-26522(2000).
 [7]
 RN
 RP VARIANT PD THR-53.
 RX MEDLINE-97342853; PubMed-9197268;
 RA Polymeropoulos M.H., Lavedan C., Leroy E., Ide S.E., Dehejia A.,
 RA Chandra A., Pike B., Root H., Rubenstein J., Boyer R., Stenroos E.S.,
 RA Chandrasekharappa S., Athanassiadou A., Papapetropoulos T.,
 RA Johnson W.G., Lazzarini A.M., Duvoisin R.C., di Iorio G., Golbe L.I.,
 RA Nussbaum R.L.;
 RA "Mutation in the alpha-synuclein gene identified in families with
 RT Parkinson's disease.";
 RL Science 276:2045-2047(1997).
 [8]
 RN
 RP VARIANT PD PRO-30.
 RX MEDLINE-98122567; PubMed-9462735;
 RA Krueger R., Kuhn W., Mueller T., Woitalla D., Graeber M., Koesel S.,
 RA Przuntek H., Epplen J.T., Schoels L., Riess O.;
 RT "A30Pro mutation in the gene encoding alpha-synuclein in Parkinson's
 disease.";
 RL Nat. Genet. 18:106-108(1998).
 CC
 CC -!- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY. IT COULD ACT AS
 CC A MOLECULAR CHAPERONE THAT MEDIATES THE TRANSFORMATION OF SOLUBLE
 CC A BETA INTO INSOLUBLE AMYLOID.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2-4 AND
 CC 2-5; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXPRESSED PRINCIPALLY IN BRAIN BUT IS ALSO
 CC EXPRESSED IN LOW CONCENTRATIONS IN ALL TISSUES EXAMINED EXCEPT IN
 CC LIVER.
 CC -!- PTM: PHOSPHORYLATED, PREDOMINANTLY ON SERINE RESIDUES.
 CC PHOSPHORYLATION BY CK1 APPEARS TO OCCUR ON RESIDUES DISTINCT FROM
 CC THE RESIDUE PHOSPHORYLATED BY OTHER KINASES.
 CC -!- DISEASE: DEFECTS IN SNCA ARE A CAUSE OF A FAMILIAL FORM OF
 CC PARKINSON'S DISEASE (FPD), A COMMON NEURODEGENERATIVE DISORDER
 CC CHARACTERIZED BY PROGRESSIVE NEURONAL DEGENERATION WHICH
 CC PREDOMINANTLY AFFECTS THE DOPAMINERGIC NERVE CELLS OF THE
 CC SUBSTANTIA NIGRA IN THE BRAIN.
 CC -!- DISEASE: DEPOSITION OF FIBRILLAR AMYLOID PROTEINS INTRANEURONALLY
 CC AS NEUROFIBRILLARY TANGLES IS CHARACTERISTIC OF ALZHEIMER'S
 CC DISEASE (AD). THIS IS A MINOR PROTEIN FOUND WITHIN THESE DEPOSITS.
 CC -!- DISEASE: BRAIN IRON ACCUMULATION TYPE 1 (NBIA1, ALSO CALLED
 CC HALLERVORDEN-SPATZ SYNDROME), A RARE NEUROAXONAL DYSTROPHY, IS
 CC HISTOLOGICALLY CHARACTERIZED BY AXONAL SPHEROIDS, IRON DEPOSITION,
 CC LEWY BODY (LB)-LIKE INTRANEURONAL INCLUSIONS, GLIAL INCLUSIONS AND
 CC NEUROFIBRILLARY TANGLES. SNCA IS FOUND IN LB-LIKE INCLUSIONS,
 CC GLIAL INCLUSIONS AND SPHEROIDS.
 CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L08850; AAA16117.1; -
 CC EMBL; L36674; AAA98493.1; -
 CC EMBL; L36675; BAA98487.1; -
 CC EMBL; D31839; BAA06625.1; -
 CC EMBL; U46901; AAC02114.1; -
 CC EMBL; U46897; AAC02114.1; JOINED.
 CC EMBL; U46898; AAC02114.1; JOINED.
 CC EMBL; U46899; AAC02114.1; JOINED.
 CC Genes; HGNC:11138; SNCA.
 CC MIM; 163890; -
 CC MIM; 601508; -
 CC InterPro; IPR001058; Synuclein.
 CC Pfam; PF01387; Synuclein; 1.
 CC PRINTS; PR01211; SYNUCLEIN.
 CC Prodom; PD010631; Synuclein; 1.
 KW Repeat; Amyloid; Alzheimer's disease; Alternative splicing;

Disease mutation; Phosphorylation.
 4 X 11 AA TANDEM REPEATS OF [EGS]-K-T-K-
 [EQ]-[GQ]-V-X(4).
 FT DOMAIN 20 67
 FT REPEAT 20 30 1.
 FT REPEAT 31 41 2.
 FT REPEAT 42 56 3 (APPROXIMATE).
 FT REPEAT 57 67 4.
 FT MOD_RES 87 87 PHOSPHORYLATION (MINOR).
 FT MOD_RES 129 129 PHOSPHORYLATION (MAJOR) (BY CK2, GRK2 AND
 GRK5).
 FT VARSPLIC 41 54 MISSING (IN ISOFORM 2-5).
 FT VARSPLIC 103 130 MISSING (IN ISOFORM 2-4).
 FT VARIANT 30 30 A -> P (IN PD).
 FT VARIANT 53 53 A -> T (IN PD).
 FT VARIANT 53 53 /FTID=VAR_007957.
 FT SEQUENCE 140 AA; 14460 MW; 6BB2F12128931663 CRC64;
 /FTID=VAR_007454.
 SQ
 Query Match 100.0%; Score 702; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 3.7e-47;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MDVFMKLSKAKGEGVVAAGTKGVAAGTKGEGVLYVSGTKGEGVHGVATVAEKT 60
 Db 1 MDVFMKLSKAKGEGVVAAGTKGVAAGTKGEGVLYVSGTKGEGVHGVATVAEKT 60
 Qy 61 EQVTNVGAVVGTGTAVAKTVEGAGSTAAATGFVKDQGLKNEEGAPQEGILEMDPVP 120
 Db 61 EQVTNVGAVVGTGTAVAKTVEGAGSTAAATGFVKDQGLKNEEGAPQEGILEMDPVP 120
 Qy 121 DNEAYEMPSEEGYQDYPEA 140
 Db 121 DNEAYEMPSEEGYQDYPEA 140
 RESULT 2
 SYUA_RAT STANDARD; PRT; 140 AA.
 AC P37377; P37378;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-synuclein.
 GN SNCA.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN
 [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE=Brain;
 RX MEDLINE-92099853; PubMed-1661825;
 RA Maroteaux L., Scheller R.H.;
 RT "The rat brain synucleins; family of proteins transiently associated
 RT with neuronal membrane.";
 RL Brain Res. Mol. Brain Res. 11:335-343(1991).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORM SYN2).
 RX MEDLINE-88316381; PubMed-3411354;
 RA Maroteaux L., Campanelli J.T., Scheller R.H.;
 RT "Synuclein: a neuron-specific protein localized to the nucleus and
 RT presynaptic nerve terminal.";
 RL J. Neurosci. 8:2804-2815(1988).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORM SYN1). R.E.;
 RA Kholodilov N.G., Neystat M., Burke R.E.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; SYN1 (SHOWN HERE), SYN2
 CC AND SYN3; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: FOUND ONLY IN BRAIN (HIPPOCAMPUS, BRAINSTEM
 CC AND CORTEX). SPECIFICALLY EXPRESSED IN NEURONAL CELL BODIES AND
 CC SYNAPSES.


```

CC -1: PTM: PHOSPHORYLATED, PREDOMINANTLY ON SERINE RESIDUES.
CC PHOSPHORYLATION BY CK1 APPEARS TO OCCUR ON RESIDUES DISTINCT FROM
CC THE RESIDUE PHOSPHORYLATED BY OTHER KINASES.
CC -1: SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S73007; AAB20688.1; -
DR EMBL; S73008; AAB20689.1; -
DR EMBL; S73009; AAB20690.1; -
DR EMBL; AF007758; AAC16026.1; -
DR PIR; A43959; A43959.
DR PIR; B43959; B43959.
DR PIR; C43959; C43959.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein; 1.
DR PRINTS; PR01211; SYNUCLEIN.
DR ProDom; PD010631; synuclein; 1.
KW Alternative splicing; Repeat; Phosphorylation.
FT DOMAIN 20 67
FT [EQ]-[GQ]-V-X(4).
FT REPEAT 20 30
FT REPEAT 31 41
FT REPEAT 42 56
FT REPEAT 57 67
FT MOD_RES 129 129
FT VARSPLIC 42 42
FT VARSPLIC 43 140
FT VARSPLIC 104 140
FT YPMGECTNHPRLIALRVKRSYREHSWRPRKLSLACVMD
FT PLPLT (IN ISOFORM SYN2).
SQ SEQUENCE 140 AA; 14515 MW; 1FFD19CD3B9E636C CRC64;
Query Match 95.4%; Score 670; DB 1; Length 140;
Best Local Similarity 95.0%; Pred. No. 1e-44;
Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVGVATVAEKT 60
QY 61 EQVTNKGAVVTGTTAAVAKTVEGAGSIAAATGFVKKQDLGKNEGAPOEGILEDMPVDP 120
Db 61 EQVTNKGAVVTGTTAAVAKTVEGAGSIAAATGFVKKQDLGKNEGAPOEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 SSEAYEMPSEEGYQDYEP 140
RESULT 3
ID SYUA_MOUSE STANDARD; PRT; 140 AA.
AC O55042; Q90UR3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component
DE of amyloid precursor) (NACP).
GN SNCA OR SYN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=98311217; PubMed=9648883;
RA Hsu L.J., Mallory M., Xia Y., Weinbergs I., Hashimoto M.,
RA Yoshimoto M., Thal L.J., Saitoh T., Masliah E.;
RT "Expression pattern of synucleins (non-Abeta component of Alzheimer's
RT disease amyloid precursor protein/alpha-synuclein) during murine
RT brain development.";
RL J. Neurochem. 71:338-344(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR;
RX MEDLINE=98264007; PubMed=9601701;
RA Hong L., Ko H.W., Gwag B.J., Joe E., Lee S., Kim Y.T., Suh Y.-H.;
RT "The cDNA cloning and ontogeny of mouse alpha-synuclein.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BL6, and SV129/OLA; TISSUE=Brain;
RA Fog J.U., Kallunki P.;
RT "Genomic cloning of the mouse alpha-synuclein and analysis of the
RT promoter.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RA Touchman J.W., Dehejia A., Chiba-Falek O., Cabin D.E., Schwartz J.R.,
RA Orrison B.M., Polymopoulos M.H., Nussbaum R.D.;
RT "Human and mouse alpha-synuclein genes: comparative genomic sequence
RT analysis and identification of a novel gene regulatory element.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1: FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
CC -1: PTM: PHOSPHORYLATED, PREDOMINANTLY ON SERINE RESIDUES (BY
CC SIMILARITY).
CC -1: SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF044672; AAC00521.1; -
DR EMBL; AF033261; AAD11254.1; -
DR EMBL; AF179273; AAD56908.1; -
DR EMBL; AF179272; AAD56907.1; -
DR EMBL; AF179268; AAD56907.1; JOINED.
DR EMBL; AF179269; AAD56907.1; JOINED.
DR EMBL; AF179270; AAD56907.1; JOINED.
DR EMBL; AF179271; AAD56907.1; JOINED.
DR EMBL; AF163865; AAG30304.1; -
DR MGB; MGI:1277151; Snca.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein; 1.
DR PRINTS; PR01211; SYNUCLEIN.
DR ProDom; PD010631; Synuclein; 1.
KW Repeat; Phosphorylation.
FT DOMAIN 20 67
FT [EQ]-[GQ]-V-X(4).
FT REPEAT 20 30
FT REPEAT 31 41
FT REPEAT 42 56
FT REPEAT 57 67
FT MOD_RES 129 129
FT CONFLICT 58 58
SQ SEQUENCE 140 AA; 14485 MW; 1FFD19D7E15E636C CRC64;
Query Match 95.3%; Score 669; DB 1; Length 140;
Best Local Similarity 95.0%; Pred. No. 1.2e-44;
Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVVGHVGVATVAEKT 60

```


Db 1 MDVPMKLSKAKGVAEAAEKTQGVAAAGKTKEGVLVYVGSKTKEGVHGVTTVAEKT 60
 QY 61 EQVTVNGGAVVGTVAQAQKTVEGAGSIAAATGFKVKKDQLGKNEEGAPQEGILEDMVPDP 120
 Db 61 EQVTVNGGAVVGTVAQAQKTVEGAGSIAAATGFKVKKDQMGKEEGYPOEGILEDMVPDP 120

QY 121 DNEAYEMPSSEGYQDYEPPEA 140
 Db 121 GSEAYEMPSSEGYQDYEPPEA 140

RESULT 4

SYUA_POEGU
 ID SYUA_POEGU STANDARD; PRT; 143 AA.
 AC Q91448;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-synuclein (Synelfin).
 OS Poephila guttata (zebra finch) (Taeniopygia guttata).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Estrilidae;
 OC Estrilidae; Taeniopygia.
 OX NCBI_TaxID=59729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=95374791; PubMed=7646890;
 RT "George J.M., Jin H., Woods W.S., Clayton D.F.;
 RT "Characterization of a novel protein regulated during the critical
 RT "period for song learning in the zebra finch.";
 RL Neuron 15:361-372(1995).
 CC -1- FUNCTION: MAY SERVE A FUNCTION CRITICAL TO THE REGULATION OF
 CC VERTEBRATE NEURAL PLASTICITY.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33860; AAA93538.1;
 CC InterPro: IPR001038; Synuclein.
 CC Pfam: PF01387; Synuclein; 1.
 CC PRINTS: PR01211; SYNUCLEIN.
 CC ProDom: PD010631; Synuclein; 1.
 KW Repeat.
 FT DOMAIN 20 67 4 X 11 AA TANDEM REPEATS OF [EGS]-K-T-K-
 FT [EQ]-[GQ]-V-X(4).
 FT REPEAT 20 30 1.
 FT REPEAT 31 41 2.
 FT REPEAT 42 56 3 (APPROXIMATE).
 FT REPEAT 57 67 4.
 FT SEQUENCE 143 AA; 14874 MW; 9C3AB9C9902C54A7 CRC64;

Query Match 84.8%; Score 595.5; DB 1; Length 143;
 Best Local Similarity 86.7%; Pred. No. 4.9e-39;
 Matches 124; Conservative 6; Mismatches 10; Indels 3; Gaps 2;

QY 1 MDVPMKLSKAKGVAEAAEKTQGVAAAGKTKEGVLVYVGSKTKEGVHGVTTVAEKT 60
 Db 1 MDVPMKLSKAKGVAEAAEKTQGVAAAGKTKEGVLVYVGSKTKEGVHGVTTVAEKT 60
 QY 61 EQVTVNGGAVVGTVAQAQKTVEGAGSIAAATGFKVKKDQLGK-NEEGAPQEGILEDMVPDP 117
 Db 61 EQVTVNGGAVVGTVAQAQKTVEGAGSIAAATGFKVKKDQLGK-NEEGAPQEGILEDMVPDP 120
 QY 118 VDPNEAYEMPSSEGYQDYEPPEA 140

Db 121 VDPNEAYEMPSSEGYQDYEPPEA 143

RESULT 5

SYUB_HUMAN
 ID SYUB_HUMAN STANDARD; PRT; 134 AA.
 AC Q16143;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-synuclein.
 GN SNCB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=94252398; PubMed=8194594;
 RA Jakes R., Spillantini M.G., Goedert M.;
 RT "Identification of two distinct synucleins from human brain.";
 RL FEBS Lett. 345:27-32(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99026142; PubMed=9806846;
 RA Lavedan C., Leroy E., Torres R., Dehejia A., Dutra A., Buchholtz S.,
 RA Nussbaum R.L., Polymeropoulos M.H.;
 RT "Genomic organization and expression of the human beta-synuclein gene
 RT (SNCB).";
 RL Genomics 54:173-175(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PHOSPHORYLATION.
 RX MEDLINE=20409007; PubMed=10852916;
 RA Pronin A.N., Morris A.J., Surguchov A., Benovic J.L.;
 RT "Synucleins are a novel class of substrates for G protein-coupled
 RT receptor kinases.";
 RL J. Biol. Chem. 275:26515-26522(2000).
 CC -1- FUNCTION: MAY BE INVOLVED IN NEURONAL PLASTICITY.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN; CONCENTRATED
 CC IN PRESYNAPTIC NERVE TERMINALS.
 CC -1- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY G-PROTEIN COUPLED RECEPTOR
 CC KINASES (GKR) IS MORE EFFICIENT THAN PHOSPHORYLATION BY CK1, CK2
 CC AND CAM-KINASE II.
 CC -1- DISEASE: BRAIN IRON ACCUMULATION TYPE 1 (NB1A), ALSO CALLED
 CC HALLERVORDEN-SPATZ SYNDROME), A RARE NEUROAXONAL DYSTROPHY, IS
 CC HISTOLOGICALLY CHARACTERIZED BY AXONAL SPHEROIDS, IRON DEPOSITION,
 CC LEWY BODY (LB)-LIKE INTRANEUROAL INCLUSIONS, GLIAL INCLUSIONS AND
 CC NEUROFIBRILLARY TANGLES. SNCB IS FOUND IN SPHEROIDS BUT NOT IN
 CC INCLUSIONS.
 CC -1- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S69965; AAB30860.1;
 CC EMBL: AF053136; AAC80286.1;
 CC EMBL: AF053134; AAC80286.1; JOINED.
 CC EMBL: AF053135; AAC80286.1; JOINED.
 CC EMBL: BC002902; AAC02902.1;
 CC Genew; HGNC:11140; SNCB.


```

RT receptor kinases."
RL J. Biol. Chem. 275:26515-26522(2000).
CC -!- FUNCTION: PLAYS A ROLE IN NEUROFILAMENT NETWORK INTEGRITY. MAY BE
CC INVOLVED IN MODULATING AXONAL ARCHITECTURE DURING DEVELOPMENT AND
CC IN THE ADULT. IN VITRO, INCREASES THE SUSCEPTIBILITY OF
CC NEUROFILAMENT-H TO CALCIUM-DEPENDENT PROTEASES (BY SIMILARITY).
CC MAY ALSO FUNCTION IN MODULATING THE KERATIN NETWORK IN SKIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, PARTICULARLY IN THE
CC SUBSTANTIA NIGRA. ALSO EXPRESSED IN THE CORPUS CALLOSUM, HEART,
CC SKELETAL MUSCLE, OVARY, TESTIS, COLON AND SPLEEN. WEAK EXPRESSION
CC IN PANCREAS, KIDNEY AND LUNG.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY GRK5 APPEARS TO OCCUR ON
CC RESIDUES DISTINCT FROM THE RESIDUE PHOSPHORYLATED BY OTHER
CC KINASES.
CC -!- DISEASE: BRAIN IRON ACCUMULATION TYPE 1 (NBIA1, ALSO CALLED
CC HALLERVORDEN-SPATZ SYNDROME), A RARE NEUROAXONAL DYSTROPHY, IS
CC HISTOLOGICALLY CHARACTERIZED BY AXONAL SPHEROIDS, IRON DEPOSITION,
CC LEWY BODY (LB)-LIKE INTRANEUROAL INCLUSIONS, GLIAL INCLUSIONS AND
CC NEUROFIBRILLARY TANGLES. SNCG IS FOUND IN SPHEROIDS BUT NOT IN
CC INCLUSIONS.
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF010126; AAB64109.1; -
CC EMBL; AF037207; AAC36586.1; -
CC EMBL; AF044311; AAC37738.1; -
CC EMBL; AF017256; AAC36550.1; -
CC EMBL; BC014098; AAH14098.1; -
CC Genew; HGNC:11141; SNCG.
CC MIM; 602998; -
CC InterPro; IPR001058; Synuclein.
CC Pfam; PF01387; Synuclein; 1.
CC PRINTS; PRO1211; Synuclein.
CC ProDom; PD010631; Synuclein; 1.
CC Repeat; Polymorphism; Phosphorylation.
CC DOMAIN 20 67
CC [EQ]-[GQ]-V-X(4).
CC
CC REPEAT 20 30
CC 1.
CC 2.
CC 3 (APPROXIMATE).
CC 4.
CC PHOSPHORYLATION (BY CK2, CAMK2 AND GRK2).
CC V->E.
CC /FTId=VAR_007455.
CC CONFLICT 13 13 E -> K (IN REF. 1).
CC CONFLICT 68 68 E -> K (IN REF. 1).
CC F19C046302DBE688 CRC64;
CC SEQUENCE 127 AA; 13301 MW; 13301 MW; F19C046302DBE688 CRC64;
CC
CC Query Match 44.6%; Score 313; DB 1; Length 127;
CC Best Local Similarity 62.8%; Pred. No. 1.5e-17;
CC Matches 71; Conservative 9; Mismatches 29; Indels 4; Gaps 2;
CC
CC QY 1 MDVFMKGLSKAEKVVAAEKTKQGVAAEAKTKGKGVLYVSKTKGKGVVGVATVAEKTK 60
CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 1 MDVFMKGFSAIEKGVVAVGAEKTKQGVAAEAKTKGKGVVGVATVAEKTK 60
CC
CC QY 61 EQVTVGGAVTVGTAVAGKTVGAGSAGSAAATGFVKKQDLGKNRCAP-QEGI 112
CC || ||||| ||||| ||||| ||||| ||||| ||||| |||||
CC Db 61 EQANAVSEAVSSVNTVATKTVAAENIAVTSGVVRKEDL---RPSAPQEGV 110
CC
CC RESULT 10
CC SYUG_MOUSE
CC ID SYUG_MOUSE STANDARD; PRT: 123 AA.

```

```

AC Q9Z0F7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-synuclein (Persyn).
GN SNCG OR PERSYN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=CD-1; TISSUE=Embryonic ganglion;
RX MEDLINE=99019751; PubMed=9801372; Thompson J., Privalova E.M.,
RA Buchman V.L., Hunter H.J., Pinon L.G., Thompson J., Privalova E.M.,
RA Ninkina N.N., Davies A.M.;
RT "Persyn, a member of the synuclein family, has a distinct pattern of
RL expression in the developing nervous system.";
RN J. Neurosci. 18:9335-9341(1998).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=CD-1;
CC MEDLINE=99162411; PubMed=10051410;
CC RA Alimova-Kost M.V., Ninkina N.N., Imreh S., Gnuchev N.V., Adu J.,
CC Davies A.M., Buchman V.L.;
CC "Genomic structure and chromosomal localization of the mouse persyn
CC gene.";
CC Genomics 56:224-227(1999).
CC [3]
CC POSSIBLE FUNCTION IN SKIN.
CC MEDLINE=99126387; PubMed=9925745;
CC RA Ninkina N.N., Privalova E.M., Pinon L.G., Davies A.M., Buchman V.L.;
CC "Developmentally regulated expression of persyn, a member of the
CC synuclein family, in skin.";
CC Exp. Cell Res. 246:308-311(1999).
CC -!- FUNCTION: PLAYS A ROLE IN NEUROFILAMENT NETWORK INTEGRITY. MAY BE
CC INVOLVED IN MODULATING AXONAL ARCHITECTURE DURING DEVELOPMENT AND
CC IN THE ADULT. IN VITRO, INCREASES THE SUSCEPTIBILITY OF
CC NEUROFILAMENT-H TO CALCIUM-DEPENDENT PROTEASES. MAY ALSO FUNCTION
CC IN MODULATING THE KERATIN NETWORK IN SKIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, PARTICULARLY IN THE
CC SUBSTANTIA NIGRA. ALSO EXPRESSED IN THE CORPUS CALLOSUM, HEART,
CC SKELETAL MUSCLE, OVARY, TESTIS, COLON AND SPLEEN. WEAK EXPRESSION
CC IN PANCREAS, KIDNEY AND LUNG. EXPRESSED PREDOMINANTLY IN THE CELL
CC BODIES AND AXONS OF PRIMARY SENSORY NEURONS, SYMPATHETIC NEURONS
CC AND MOTONEURONS.
CC -!- DEVELOPMENTAL STAGE: DEVELOPMENTALLY EXPRESSED IN PRIMARY SENSORY
CC NEURONS AND MOTONEURONS. IN TRIGEMINAL GANGLIA, EXPRESSION
CC INCREASES BETWEEN EMBRYONIC DAY 10 AND DAY 12. HIGH LEVELS ARE
CC MAINTAINED HERE THROUGHOUT LATER STAGES OF DEVELOPMENT AND IN
CC ADULTHOOD.
CC -!- PTM: PHOSPHORYLATED. PHOSPHORYLATION BY GRK5 APPEARS TO OCCUR ON
CC RESIDUES DISTINCT FROM THE RESIDUE PHOSPHORYLATED BY OTHER
CC KINASES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SYNUCLEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF017255; AAC98893.1; -
CC EMBL; AF099986; AAD19899.1; -
CC EMBL; AF099984; AAD19899.1; JOINED.
CC EMBL; AF099985; AAD19899.1; JOINED.
CC SWISS-2DPAGE; Q9Z0F7; MOUSE.
CC MGI; MGI:1298397; SnCG.
CC InterPro; IPR001058; Synuclein.
CC Pfam; PF01387; Synuclein; 1.

```


Search completed: January 28, 2003, 14:15:02
Job time : 4.14149 secs


```

Db 61 EQVNVGGAVVGTGTAQAQKTVGAGNIAAATGLVKKDQLAKONERGFLEQGMVNTDIP 120
Qy 118 VDPNEAYEMPSBEGQDYRPEA 140
   |||:||||||| || |||||
Db 121 VDPENEAYEMPEEYQDYRPEA 143

RESULT 2
Q9CXF8 ID Q9CXF8 PRELIMINARY; PRT; 122 AA.
AC Q9CXF8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 14 days embryo liver cDNA, RIKEN full-length enriched library,
DE clone:4430402N11, full insert sequence.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC LIVER;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Willing L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014472; BAB29375.1; -.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein; 1.
DR PRINTS; PR01211; SYNUCLEIN.
DR ProDom; PD010631; Synuclein; 1.
SQ SEQUENCE 122 AA; 12432 MW; 2D03AFC4944C78D CRC64;

Query Match 69.1%; Score 485; DB 11; Length 122;
Best Local Similarity 97.1%; Pred. No. 2e-30;
Matches 99; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLVVGSKTKEGVVHGVATVAEKTK 60
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLVVGSKTKEGVVHGVATVAEKTK 60
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 EQVTNVGGAVVGTGTAQAQKTVGAGSIAAATGFVKKDQLG 102
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 EQVTNVGGAVVGTGTAQAQKTVGAGSIAAATGFVKKDQLG 102
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 EQVTNVGGAVVGTGTAQAQKTVGAGSIAAATGFVKKDQLG 102
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 EQVTNVGGAVVGTGTAQAQKTVGAGSIAAATGFVKKDQLG 102
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
Q91Z23 ID Q91Z23 PRELIMINARY; PRT; 133 AA.
AC Q91Z23;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-synuclein (Synuclein, beta).
GN A1838531 OR SNCB.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=21367613; PubMed=11474193;
RA Sopher B.L., Koszidin K.L., McClain M.E., Myrick S.B., Martinez R.A.,
RA Smith A.C., La Spada A.R.;
RT "Genomic organization, chromosome location, and expression analysis of
RT mouse beta-synuclein, a candidate for involvement in
RT neurodegeneration.";
RL Cytogenet. Cell Genet. 93:117-123(2001).
RN 121
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF348164; AAK83238.1; -.
DR EMBL; AF348162; AAK83238.1; JOINED.
DR EMBL; AF348163; AAK83238.1; JOINED.
DR EMBL; BC019409; AAK83238.1; JOINED.
DR MGD; MGI:1889011; Sncb.
DR MGD; MGI:2138328; A1838531.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein; 1.
DR ProDom; PD010631; Synuclein; 1.
SQ SEQUENCE 133 AA; 14052 MW; 8274D8A6A0D8E4D5 CRC64;

Query Match 55.1%; Score 386.5; DB 11; Length 133;
Best Local Similarity 61.0%; Pred. No. 9.5e-23;
Matches 89; Conservative 18; Mismatches 20; Indels 19; Gaps 5;

Qy 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLVVGSKTKEGVVHGVATVAEKTK 60
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAEAGTKEGVLVVGSKTKEGVVHGVATVAEKTK 59
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 61 EQVTNVGGAVVGTGTAQAQKTVGAGSIAAATGFVKKDQLG 114
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 60 EQASHLGGAVFS-----GAGNIAAATGLVKKKEEFTDLKPEVAQEAEEPLIE 108
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 115 DMPVDPDNEAYEMPSBEGQDYRPEA 140
   |||:||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109 PL-MEPEGESYEDSPQEEYQDYRPEA 133

RESULT 4
Q919G9 ID Q919G9 PRELIMINARY; PRT; 133 AA.
AC Q919G9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta-synuclein.
GN SNCB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=20544973; PubMed=11091093;
RA Tiunova A.A., Anokhin K.V., Saha A.R., Schmidt O., Hanger D.P.,
RA Anderton B.H., Davies A.M., Ninkina N.N., Buchman V.L.;
RT "Chicken synuclein: cloning and expression in the developing
RT embryo.";
RL Mech. Dev. 99:195-198(2000).
DR EMBL; AF253514; AAF67730.1; -.
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein; 1.
DR PRINTS; PR01211; SYNUCLEIN.
DR ProDom; PD010631; Synuclein; 1.

```



```

SQ SEQUENCE 133 AA; 14063 MW; 8F84E1A9EDC3C6F2 CRC64;
Query Match
Best Local Similarity 53.1%; Score 372.5; DB 13; Length 133;
Matches 87; Conservative 17; Mismatches 23; Indels 19; Gaps 5;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVLVGVATVAEAKTK 60
D 1 MEVFMKGLSKAKGVVAAAEKTKQGVAAAEKTKGVLVVGSKTKGVLVGVATVAEAKTK 59
QY 61 EQVTNVGGAVVTGVTAVAKTVEGAGSIAAATGFVKKDKQLG---KNEGAPQEGILEDMP 117
D 60 EQASQGEAFAFS-----GAGNIAAATGLVKKEFPADLKAEEVA-QEAVEEPLV 107
QY 118 ---VDPNEAYEMPSEEGYQDYEPEA 140
D 108 EPLLEPEGESYEESPOEEYQVEPEA 133

RESULT 5
Q90WV3 PRELIMINARY; PRT; 129 AA.
AC Q90WV3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Synuclein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP "Developmental expression of Xenopus synuclein.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY055119; AAL16411.1;
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein.1.
DR ProDom; PD010631; Synuclein.1.
SQ SEQUENCE 129 AA; 13442 MW; 2CD1B7B067AC708B CRC64;

Query Match
Best Local Similarity 51.8%; Score 363.5; DB 13; Length 129;
Matches 79; Conservative 13; Mismatches 30; Indels 15; Gaps 1;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVLVGVATVAEAKTK 60
D 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAEKTKGVLVVGSKTKGVLVGVATVAEAKTK 60
QY 61 EQVTNVGGAVVTGVTAVAKTVEGAGSIAAATGFVKKDKQLGKNEGAPQEGILEDMPVDP 120
D 61 EQANVVGAVSGVNVQVSKTVEGTENVVSSGLVKKEDL-----HPDQP 105
QY 121 DNEAYEMPSEEGYQDYE 137
D 106 EEPAAEPAVEATESTE 122

RESULT 6
Q919H0 PRELIMINARY; PRT; 128 AA.
AC Q919H0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Persyn.
GN SNG1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.

```

```

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20544973; PubMed=11091093;
RA Tlunova A.A., Anokhin K.V., Saha A.R., Schmidt O., Hanger D.P.,
RA Antonov B.H., Davies A.M., Ninkina N.N., Buchman V.L.;
RT "Chicken synucleins: cloning and expression in the developing
RT embryo.";
RL Mech. Dev. 99:195-198(2000).
DR EMBL; AF253513; AAF67729.1;
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein.1.
DR PRINTS; PR01211; SYNNUCLEIN.
DR ProDom; PD010631; Synuclein.1.
SQ SEQUENCE 128 AA; 12966 MW; A5B944446053A2DA7 CRC64;

Query Match
Best Local Similarity 49.6%; Score 348.5; DB 13; Length 128;
Matches 84; Conservative 7; Mismatches 35; Indels 13; Gaps 2;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVLVGVATVAEAKTK 60
D 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAEKTKGVLVVGSKTKGVLVGVATVAEAKTK 60
QY 61 EQVTNVGGAVVTGVTAVAKTVEGAGSIAAATGFVKKDKQLGKNEGAPQEGILEDMPVDP 120
D 61 EQANVVGAVSVNTVANKTVEGAGTIVATGVVKEDL-----APQQAPE----- 108
QY 121 DNEAYEMPSEEGYQDYEPE 139
D 109 -GEAIPGSTEGGEGENE 126

RESULT 7
Q96P61 PRELIMINARY; PRT; 127 AA.
AC Q96P61;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Synuclein gamma.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han C., Zhang B., Peng X., Yuan J., Qiang B.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411524; AAL05870.1;
DR InterPro; IPR001058; Synuclein.
DR Pfam; PF01387; Synuclein.1.
DR ProDom; PD010631; Synuclein.1.
SQ SEQUENCE 127 AA; 13389 MW; D758B345DBFF7578 CRC64;

Query Match
Best Local Similarity 44.1%; Score 309.5; DB 4; Length 127;
Matches 69; Conservative 9; Mismatches 29; Indels 3; Gaps 1;

QY 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAGTKGVLVVGSKTKGVLVGVATVAEAKTK 60
D 1 MDVFMKGLSKAKGVVAAAEKTKQGVAAAEKTKGVLVVGSKTKGVLVGVATVAEAKTK 60
QY 61 EQVTNVGGAVVTGVTAVAKTVEGAGSIAAATGFVKKDKQLGKNEGAPQEGILEDMP 110
D 61 EQANVSEAVSVSVNTVATKVEEAENIAVTSGVVRKEDL---RPSAPQ 107

RESULT 8
Q9EQC3 PRELIMINARY; PRT; 54 AA.
ID Q9EQC3
AC Q9EQC3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```


ID	Q95ZV0	PRELIMINARY:	PRT;	750 AA.
AC	Q95ZV0;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	F25H8.5c protein.			
GN	F25H8.5C.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;			
OC	Rhabditiidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RA	Gajadsty S.;			
RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology.";			
RL	Science 282:2012-2018(1998).			
DR	EMBL; Z69360; CAC42291.1; -.			
DR	InterPro; IPR004238; LEA.			
DR	Pfam; PF02987; LEA; 5.			
SQ	SEQUENCE 750 AA; 77336 MW; A68CA612D8345281 CRC64;			
	Query Match 17.4%; Score 122; DB 5; Length 750;			
	Best Local Similarity 32.7%; Pred. No. 0.24; Indels 20; Gaps			
	Matches 35; Conservative 13; Mismatches 39;			
QY	13 EGVVAAAEKTKQCVAEAGKTKEGVLYYSKSTKEGVHVGVATVAETK----EQVTNVGA 69			
DB	: : : : :			
	610 DGAVAKDAAVEGATTAGNKTVEGRFVG----ENVAHGAGKVADGAKAAGDTVVEGKA 66			
QY	70 VW-----TGYTVAQKTVEGAGSIAAATGFVKKDLGLNGESGAPOEGI 112			
DB	: :			
	666 AAETESGIAFVGOKIABGAGAVAGG-----AKAAGGAVVDGV 703			
RESULT 11				
Q19790	Q19790	PRELIMINARY:	PRT;	798 AA.
ID	Q19790;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	F25H8.5 protein.			
GN	F25H8.5.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;			
OC	Rhabditiidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID=6239;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RA	Gajadsty S.;			
RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.			
CC	- - ALTERNATIVE PRODUCTS: TWO FORMS (A AND B) MAY BE PRODUCED BY			
CC	OF FORM A.			
CC	EMBL; Z69360; CAA93285.1; -.			
DR	EMBL; Z69360; CAA93286.1; -.			
DR	InterPro; IPR004238; LEA.			
DR	InterPro; IPR002965; P_rich_extensn.			
DR	Pfam; PF02987; LEA; 6.			
DR	PRINTS; PR01217; PRICHEXTENS.			
KW	Alternative splicing; Hypothetical protein.			
FT	VARSPLIC 387 387 N -> S (IN ISOFORM B).			
FT	VARSPLIC 388 798 MISSING (IN ISOFORM B).			
SQ	SEQUENCE 798 AA; 81507 MW; 4A61S20E2313BF2 CRC64;			
	Query Match 17.4%; Score 122; DB 5; Length 798;			
	Best Local Similarity 32.7%; Pred. No. 0.25;			

Matches 35; Conservative 13; Mismatches 39; Indels 20; Gaps 4;

QY 13 EGVAATAEKTGCGVAAAGKTKGVLVVGSKTKGVVHGVATVAEKT---EQVTVNGCA 69
 Db 610 DGAVAAKDAVEGATTAGNKTGEMRFVG----ENVHAGAGKADGAKAAGDTVVVEGGKA 665

QY 70 VV-----TGVTVAQAQKTVGAGSATAATGFKKDLGKNEEGAPQEG 112
 Db 666 AAEKTESGIAFGQKLAEGAGAVAGG-----AKAAGGAVVDGV 703

RESULT 12

Q39846
 ID Q39846 PRELIMINARY; PRT; 643 AA.
 AC Q39846;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE LEA protein.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SHI-SHI; TISSUE=COTYLEDON;
 RX MEDLINE=98418627; PubMed=9747855;
 RA Hsing Y.C., Tsou C.H., Hsu T.F., Chen Z.Y., Hsieh K.L., Hsieh J.S.,
 RA Chow T.Y.;
 RT "Tissue- and stage-specific expression of a soybean (Glycine max L.)
 RT seed-maturation, biotinylated protein.";
 RL Plant Mol. Biol. 38:481-490(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SHI-SHI; TISSUE=COTYLEDON;
 RA Shu T.F., Hsieh K.L., Hsing Y.I., Chen Z.Y., Chow T.Y.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U59626; AAC61783.1; -.
 SQ SEQUENCE 643 AA; 67878 MW; 0D1863D0D4196985 CRC64;

Query Match 15.0%; Score 105; DB 10; Length 643;
 Best Local Similarity 28.2%; Pred. No. 4;
 Matches 48; Conservative 10; Mismatches 68; Indels 44; Gaps 7;

QY 2 DVFMGLSKAKGEGVAAAEKTKQGVAAEAGTKGEG----VLVYVSK---TKGTVHG--- 51
 Db 289 DTISSNAKTASEKTAPVAEAKADYTLQAAEAKKASAGGTTASYVGBKAVQAKDVAVESGKS 348

QY 52 ----VATVAEKTKEQVTVNGGA-----VVTGVTVAQAQKTVGAG----- 86
 Db 349 AAGYAAKVAADLRDKATATVGMWAAHFAEKTVEGTAKAAH-VVEGAAGYAGHKAELASM 407

QY 87 -----STAAATGFKVKDLGKNEEGAPQEGILEMPPVDPDNEAYEMPSE 130
 Db 408 SAGAVKGLAASAGETAKEYTAKKKEAQRE--LEAKKPSQPOEAERPSE 455

RESULT 13

Q03967
 ID Q03967 PRELIMINARY; PRT; 317 AA.
 AC Q03967;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Group 3 late embryogenesis abundant protein (fragment).
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.

TISSUE=EMBRYONIC AXES;
 MEDLINE=93222487; PubMed=8467082;
 RA Walker-Simmons M.K., Curry J.;
 RT "Unusual sequence of group 3 LEA (II) mRNA inducible by dehydration
 RT stress in wheat.";
 RL Plant Mol. Biol. 21:907-912(1993).
 DR EMBL; M72395; AAA34267.1; -.
 DR InterPro; IPR004238; LEA.
 DR Pfam; PF02987; LEA; 3.
 FT NON_TER 1
 SQ SEQUENCE 317 AA; 33367 MW; 75BB0F431220160C CRC64;

Query Match 14.2%; Score 100; DB 10; Length 317;
 Best Local Similarity 29.6%; Pred. No. 4.1;
 Matches 53; Conservative 10; Mismatches 62; Indels 54; Gaps 8;

QY 6 KGLSKAKGEGVVAEAEKTKQGVAAE-----AAGTKGVLVVGSKTKGVVHGVATVAE 57
 Db 9 EGKEGASEMADAAAGKTKDAAAEKTRMADAAAGKTKETKDAAVEKTRE-MADTAATKAA 67

QY 58 KTK-----EQVT-----NVGAVVT-----GVTAVAQKTVGA--- 85
 Db 68 ETKDAAAEKASGAGEMVTEKARSARDAADAKASGAETVTEKAKGAKDAALDTAEGAKEY 127

QY 86 -----GSIAATGFKVKDLGKNEEGAPQEGILEMPPVDPDNEAYEMPSEGVQDVEP 138
 Db 128 MVDKKEDARRALAGSV-KDSKGETNESACQOG-----QDVRRRAAEKAEARQTHOP 179

RESULT 14

Q8SVY9
 ID Q8SVY9 PRELIMINARY; PRT; 166 AA.
 AC Q8SVY9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Hypothetical protein ECU03_1610.
 GN ECU03_1610.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornilliot E., Metenier G., Thomarat F.,
 RA Prentier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Welissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 RT Encephalitozoon cuniculi.";
 RL Nature 414:450-453(2001).
 DR EMBL; AL590443; CAD26304.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 166 AA; 18030 MW; C1E7E4EA2E31D181 CRC64;

Query Match 14.1%; Score 99; DB 5; Length 166;
 Best Local Similarity 30.5%; Pred. No. 2.2;
 Matches 39; Conservative 11; Mismatches 62; Indels 16; Gaps 4;

QY 10 KAKEGVVAAAEKTKQGVAAEAGKTKGVLVVGSKTKGVVHGVATVAEKTKEQVTVNGGA 69
 Db 21 KANDKACEARDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGS 72

QY 70 VVTGVTVAQAQKTVGAGSATAATGFKKDLGKNEEGAPQEGILEMPPVDPD-NEA 124
 Db 73 AKDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGSAKDKTKETAGS 129

QY 125 YEMPSEEG 132
 Db 130 YENKEGG 137

RESULT 15

Q971D3 PRELIMINARY: PRT: 380 AA.
 AC Q971D3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein ST1420.
 GN ST1420.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 Crenarchaeon, Sulfolobus tokodaii strain 7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000986; BAB66487.1; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 380 AA: 42015 MW: 859AC76EE032D719 CRC64;

Query Match 14.1%; Score 99; DB 17; Length 380;
 Best Local Similarity 28.6%; Pred. No. 6.1;
 Matches 42; Conservative 19; Mismatches 66; Indels 20; Gaps 6;
 QY 2 DVFEMKGLS-----KKEGVVAAREKTKQGVAAAGKTK-EGVLYVGSKTKEGVVHGV 52
 Db 197 DVFELGIPVGDLSGLPLVASRFLVNISEKWSPREMTAGETEFEGRRLLIVLKA-EGPMATV 255
 QY 53 ATVAERTKEQVTWVGAVTGTVAQKTVEG--AGSIAAATGFVKKQDLGKNEGAPQE 110
 Db 256 GTFGEAVENVVNLGCKVSRITVDAAAKLEGGQTGCSIAEGTGVANGDP-GPEKTAIERV 314
 QY 111 GILEDMPVDP-----DNEAYEMPSE 130
 Db 315 AVRYNIPIDALIVKMSMEEAITEMPRE 341

Search completed: January 28, 2003, 14:18:18
 Job time : 9.07457 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 7.80115 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKGVVAANE.....DNEAVEMPSEGVQDYEPFA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	100.0	140	16	AAR70127
2	702	100.0	140	20	AAV07277
3	702	100.0	140	20	AAW88131
4	702	100.0	140	21	AAV92093
5	702	100.0	140	21	AAV51529
6	702	100.0	140	23	ABH1545
7	697	99.3	140	21	AAV92094
8	695	99.0	140	21	AAV92090
9	693	98.7	140	21	AAV92095
10	692	98.6	140	20	AAW88132

11	692	98.6	140	21	AAV92091	Artificial alpha-s
12	692	98.6	140	21	AAV92096	Artificial alpha-sy
13	684	97.4	140	21	AAV92092	Artificial alpha-s
14	670	95.4	140	20	AAV07274	Rat synuclein alte
15	663	94.4	140	21	AAV92220	NACP/alpha-synucle
16	595.5	84.8	143	20	AAV07280	Zebrafinch (sic) s
17	485	69.1	149	20	AAV07275	Rat synuclein alte
18	427.5	60.9	133	20	AAV07279	Bovine synuclein.
19	409	58.3	134	20	AAV07278	Human beta-synucle
20	312.5	44.6	127	20	AAV07271	Human beta-syn prote
21	312.5	44.5	123	20	AAV07272	Mouse persyn prote
22	311.5	44.4	163	21	AAV58881	Breast and ovarian
23	311.5	44.4	164	22	AAO13847	Human polypeptide
24	307.5	43.8	127	19	AAW63123	Human breast cance
25	307.5	43.8	127	21	AAV87779	Human HBGBA67A pro
26	273.5	39.0	123	20	AAV07281	Rat synuclein-like
27	261	37.2	113	20	AAV07273	Torpedo californic
28	251	35.8	51	21	AAV02343	Human secreted pro
29	226	32.2	47	22	ABH43301	Peptide #10807 enc
30	226	32.2	47	22	AAW64209	Human brain expres
31	226	32.2	47	22	AAW77035	Human bone marrow
32	216	30.8	59	21	AAV02228	Human secreted pro
33	199	28.3	42	20	AAV07276	Rat synuclein alte
34	176	25.1	37	22	AAE08570	Human alpha-synuc
35	167	23.8	37	22	AAE08564	Human Beta-synuc
36	166	23.6	35	22	AAV91804	Anyloid beta-prote
37	166	23.6	35	22	AAV91831	Anyloid beta-prote
38	163	23.2	35	16	AAV70128	Novel amyloid comp
39	151	21.5	28	22	AAV65485	Human brain expres
40	151	21.5	28	22	AAV78149	Human bone marrow
41	151	21.5	28	23	ABG47162	Human peptide enco
42	143	20.4	30	22	AAV41757	Peptide #9283 enco
43	143	20.4	30	22	AAV75444	Human bone marrow
44	143	20.4	30	22	AAV35555	Peptide #9592 enco
45	142	20.2	37	22	AAE08563	Human Beta-synucle

ALIGNMENTS

RESULT 1
AAR70127
ID AAR70127 standard; Protein; 140 AA.
XX
AC AAR70127;
XX
DT 05-OCT-1995 (first entry)
XX
DE Precursor of novel amyloid component (NACP).
XX
KW Amyloid component; NAC; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT Peptide
FT Peptide
XX
PN WO9506407-A.
XX
PD 09-MAR-1995.
XX
PF 29-AUG-1994; 94WO-US09789.
XX
PR 30-AUG-1993; 93US-0114393.
XX
PA (SAIT/) SAITOH T.
XX
PI (REGC) UNIV CALIFORNIA.
XX
Masliah E;

DR WPI; 1995-115208/15.
 DR N-PSDB; AAQ83201.
 XX
 XX Novel component of amyloid in Alzheimer's disease - used in
 PT methods to diagnose and treat neuronal abnormalities in
 PT particular the deposition of amyloid plaques characteristic of AD
 XX
 XX Claim 66; Fig 2A-1 - 2A-2; 128pp; English.
 XX
 CC AAQ83201 encodes a precursor of the novel amyloid component (NACP)
 CC (AAQ70127) with an apparent mol. mass of 19,000 Da. NACP is the
 CC precursor of NAC (AAQ70128). In vivo, NACP is recovered in the
 CC cytosolic fraction of human brain homogenate. NACP has seven
 CC repeated KTRGV AA motifs, but no signal peptide sequence nor
 CC N-linked glycosylation sites. NAC is located in the most hydrophobic
 CC portion of NACP. NAC has a mol. wt. of approx. 3,500 Mr. within NAC,
 CC two new amyloid sequences have been identified known as the 'X' and
 CC 'Y' peptides which are encoded contiguously in the most hydrophobic
 CC domain (see AAQ70127 FT). The X and Y peptides are claimed (claims 68
 CC and 69 respectively).
 XX
 XX Sequence 140 AA;
 SQ
 Query Match 100.0%; Score 702; DB 16; Length 140;
 Best Local Similarity 100.0%; Pred. No. 4e-60;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGTVHGVATVAEKT 60
 DB 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGTVHGVATVAEKT 60
 QY 61 EQVTNVGGAVVTGTVTAQAQTVGAGSIAAATGFVKKDKLQKNEEGAPQEGILEMPVDP 120
 DB 61 EQVTNVGGAVVTGTVTAQAQTVGAGSIAAATGFVKKDKLQKNEEGAPQEGILEMPVDP 120
 QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140
 RESULT 2
 AAQ70277
 ID AAQ70277 standard; Protein: 140 AA.
 XX
 AC AAQ70277;
 XX
 DT 06-JUL-1999 (first entry)
 XX
 DE Human alpha-synuclein.
 XX
 KW Human; mouse; synuclein; persyn; diagnosis; neurodegenerative disorder;
 KW cancer; breast; skin; intermediate filament damage.
 XX
 OS Homo sapiens.
 XX
 PN EP908727-A1.
 XX
 PD 14-APR-1999.
 XX
 PF 21-SEP-1998; 98EP-0307628.
 XX
 PR 19-SEP-1997; 97GB-0019879.
 XX
 PA (NEUR-) EUROPA LTD.
 PA (UYSA-) UNIV ST ANDREWS.
 XX
 XX WPI; 1999-217169/19.
 DR
 XX New synuclein protein (persyn) and gene, useful in assays for
 PT screening, diagnosing or monitoring cancer, neurodegenerative
 PT disorders or skin disorders
 XX
 PS Disclosure; Page 22-23; 39pp; English.

XX This sequence represents the sequence of the human alpha-synuclein.
 CC The invention relates to novel human and mouse synuclein family members
 CC designated persyn (AAQ70271 and AAQ70172). The persyn sequence is
 CC useful for screening, diagnosing or monitoring cancer (especially breast
 CC or skin cancer), neurodegenerative disorders or skin disorders and for
 CC identifying cells having intermediate filament damage.
 XX
 XX Sequence 140 AA;
 SQ
 Query Match 100.0%; Score 702; DB 20; Length 140;
 Best Local Similarity 100.0%; Pred. No. 4e-60;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGTVHGVATVAEKT 60
 DB 1 MDVFMKGLSKAKGVAAGKTKQGVAAAGTKGVLVVGSKTKGTVHGVATVAEKT 60
 QY 61 EQVTNVGGAVVTGTVTAQAQTVGAGSIAAATGFVKKDKLQKNEEGAPQEGILEMPVDP 120
 DB 61 EQVTNVGGAVVTGTVTAQAQTVGAGSIAAATGFVKKDKLQKNEEGAPQEGILEMPVDP 120
 QY 121 DNEAYEMPSEEGYQDYEP 140
 DB 121 DNEAYEMPSEEGYQDYEP 140
 RESULT 3
 AAQ88131
 ID AAQ88131 standard; Protein: 140 AA.
 XX
 AC AAQ88131;
 XX
 DT 24-MAY-1999 (first entry)
 XX
 DE Human alpha synuclein protein.
 XX
 KW Synuclein; Parkinson disease; drug screening; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN W09859050-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 25-JUN-1998; 98WO-US13071.
 XX
 PR 25-JUN-1997; 97US-0050684.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Duvoisin RC, Johnson WG, Lavedan C, Leroy E, Nussbaum RL;
 PI Polymeropoulos MH;
 XX
 DR WPI; 1999-105624/09.
 XX
 XX New gene mutation associated with Parkinson's disease - comprising a
 PT mutation in the synuclein gene, used to develop products for
 PT detecting a predisposition to or treating Parkinson's disease
 XX
 PS Disclosure; Page 50-51; 96pp; English.
 XX
 CC A mutated human synuclein protein or homologue, associated with
 CC predisposition to Parkinson's disease (PD). The products can be used for
 CC testing for predisposition to PD. They can also be used for studying the
 CC pathophysiology of PD. They can also be used for identifying compounds
 CC for inhibiting the self-aggregation of mutant proteins that is thought
 CC to lead to PD. The products can also be used in methods directed at the
 CC correction or suppression of PD.
 XX
 XX Sequence 140 AA;
 SQ
 Query Match 100.0%; Score 702; DB 20; Length 140;


```

CC 32.
XX
SQ Sequence 140 AA;
Query Match 99.3%; Score 697; DB 21; Length 140;
Best Local Similarity 99.3%; Pred. No. 1.2e-59;
Matches 139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVAEAAEKTQGVAAEAGTKQGVLYVGSKTKEGVVHGVAETVAEKT 60
DB 1 MDVFMKGLSKAKGVAEAAEKTQGVAAEAGTKQGVLYVGSKTKEGVVHGVAETVAEKT 60
QY 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDDLGKNEEAGAPQEGILEDMPVDP 120
DB 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDDLGKNEEAGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140
RESULT 8
AAY92090
ID AAY92090 standard; Protein; 140 AA.
XX
AC AAY92090;
XX
DT 01-AUG-2000 (first entry)
XX
DE Artificial alpha-synuclein mutant B83Q+A90V.
XX
KW Artificial; alpha-synuclein mutant; accelerated aggregation;
KW animal model; inhibitor; Parkinson's Disease.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200018917-A2.
XX
PD 06-APR-2000.
XX
PF 27-SEP-1999; 99WO-US22302.
XX
PR 25-SEP-1998; 98US-0101862.
PR 24-SEP-1999; 99US-0405035.
XX
PA (AMGE-) AMGEN INC.
XX
PI Biere AL, Citron M;
XX
DR WPI; 2000-293156/25.
XX
PT New artificial mutant alpha-synuclein polypeptides are useful to
PT generate an animal model for Parkinson's disease and aggregate faster
PT than the naturally occurring alpha-synuclein and alpha-synuclein
PT mutants
XX
PS Claim 3; Page 32-33; 34pp; English.
XX
CC AAY92090-92 are artificial alpha-synuclein mutant polypeptides which
CC demonstrate accelerated aggregation compared to naturally occurring
CC alpha-synuclein mutant polypeptides. Aggregate formation was enhanced
CC for the E83Q+A90V and H50Y+A53T mutants, even when compared to the A53T
CC mutant. The aggregate formation was significantly enhanced for the
CC H50Y+A53T+A76T mutant. An in vitro aggregation assay for evaluating
CC alpha-synuclein mutants which demonstrate accelerated aggregation is
CC also claimed. Additionally, a transgenic non-human mammal which has
CC been genetically engineered to express the mutant alpha-synuclein
CC polypeptide in the brain and an in vitro aggregation assay for high
CC throughput screening of alpha-synuclein nucleation inhibitors are
CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
CC ameliorate Parkinson's Disease or related disorders (claimed). The
CC transgenic animal is used as an animal model for Parkinson's Disease.

```

```

XX
SQ Sequence 140 AA;
Query Match 99.0%; Score 695; DB 21; Length 140;
Best Local Similarity 98.6%; Pred. No. 1.9e-59;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVFMKGLSKAKGVAEAAEKTQGVAAEAGTKQGVLYVGSKTKEGVVHGVAETVAEKT 60
DB 1 MDVFMKGLSKAKGVAEAAEKTQGVAAEAGTKQGVLYVGSKTKEGVVHGVAETVAEKT 60
QY 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDDLGKNEEAGAPQEGILEDMPVDP 120
DB 61 EQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDDLGKNEEAGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
DB 121 DNEAYEMPSEEGYQDYEP 140
RESULT 9
AAY92095
ID AAY92095 standard; Protein; 140 AA.
XX
AC AAY92095;
XX
DT 01-AUG-2000 (first entry)
XX
DE Artificial alpha-synuclein mutant A53T+A30P.
XX
KW Artificial; alpha-synuclein mutant; accelerated aggregation;
KW animal model; inhibitor; Parkinson's Disease.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200018917-A2.
XX
PD 06-APR-2000.
XX
PF 27-SEP-1999; 99WO-US22302.
XX
PR 25-SEP-1998; 98US-0101862.
PR 24-SEP-1999; 99US-0405035.
XX
PA (AMGE-) AMGEN INC.
XX
PI Biere AL, Citron M;
XX
DR WPI; 2000-293156/25.
XX
PT New artificial mutant alpha-synuclein polypeptides are useful to
PT generate an animal model for Parkinson's disease and aggregate faster
PT than the naturally occurring alpha-synuclein and alpha-synuclein
PT mutants
XX
PS Example 1; Page -; 34pp; English.
XX
CC This is an artificial alpha-synuclein double mutant polypeptide
CC A53T+A30P. AAY92090-92 are artificial alpha-synuclein mutant polypeptides
CC which demonstrate accelerated aggregation compared to naturally
CC occurring alpha-synuclein mutant polypeptides. Aggregate formation was
CC enhanced for the E83Q+A90V and H50Y+A53T mutants, even when compared to
CC the A53T mutant. The aggregate formation was significantly enhanced for
CC the H50Y+A53T+A76T mutant. An in vitro aggregation assay for evaluating
CC alpha-synuclein mutants which demonstrate accelerated aggregation is
CC also claimed. Additionally, a transgenic non-human mammal which has
CC been genetically engineered to express the mutant alpha-synuclein
CC polypeptide in the brain and an in vitro aggregation assay for high
CC throughput screening of alpha-synuclein nucleation inhibitors are
CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
CC ameliorate Parkinson's Disease or related disorders (claimed). The
CC transgenic animal is used as an animal model for Parkinson's Disease.

```


CC This sequence is not given in the specification, it was constructed
 CC using the wild type sequence, which is given in the specification on page
 CC 32.

XX Sequence 140 AA; Query Match 98.7%; Score 693; DB 21; Length 140;
 XX Best Local Similarity 98.6%; Pred. No. 2.9e-59;
 XX Matches 138; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVYVSGTKGEGVHGVTVAEKT 60
 Db 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVYVSGTKGEGVHGVTVAEKT 60

QY 61 EQVTNNGGAVVTGVTVAQAQTVGAGSIAAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
 Db 61 EQVTNNGGAVVTGVTVAQAQTVGAGSIAAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYQDYEP 140
 Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 10
 AAW88132
 ID AAW88132 standard; Protein; 140 AA.

XX AC AAW88132;
 XX DT 24-MAY-1999 (first entry)
 XX DE Human alpha synuclein protein mutant.

XX KW Synuclein; Parkinson disease; drug screening; diagnosis.
 XX OS Homo sapiens.
 XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT Modified-site 53 /label= A53-D53

XX PN W09859050-A1.
 XX PD 30-DEC-1998.

XX PF 25-JUN-1998; 98WO-US13071.
 XX PR 25-JUN-1997; 97US-0050684.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PA Duvoisin RC, Johnson WG, Lavedan C, Leroy E, Nussbaum RL;
 XX PI Polymeropoulos MH;

XX WPI; 1999-105624/09.
 XX PT New gene mutation associated with Parkinson's disease - comprising a
 XX PT mutation in the synuclein gene, used to develop products for
 XX PT detecting a predisposition to or treating Parkinson's disease

XX PS Disclosure; Page-; 96pp; English.
 XX CC A mutated human synuclein protein or homologue, associated with
 XX CC predisposition to Parkinson's disease (PD). The products can be used for
 XX CC testing for predisposition to PD. They can also be used for studying the
 XX CC pathophysiology of PD. They can also be used for identifying compounds
 XX CC for inhibiting the self-aggregation of mutant proteins that is thought
 XX CC to lead to PD. The products can also be used in methods directed at the
 XX CC correction or suppression of PD.
 XX CC Note: this sequence was created from information provided in the
 XX CC specification by the indexer.

SQ Sequence 140 AA;

Query Match 98.6%; Score 692; DB 20; Length 140;
 Best Local Similarity 98.6%; Pred. No. 3.7e-59;
 Matches 138; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVYVSGTKGEGVHGVTVAEKT 60
 Db 1 MDVFMKGLSKAKGEGVVAAGKTKQGVAAAGKTKGVLVYVSGTKGEGVHGVTVAEKT 60

QY 61 EQVTNNGGAVVTGVTVAQAQTVGAGSIAAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
 Db 61 EQVTNNGGAVVTGVTVAQAQTVGAGSIAAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYQDYEP 140
 Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 11
 AAY92091
 ID AAY92091 standard; Protein; 140 AA.

XX AC AAY92091;
 XX DT 01-AUG-2000 (first entry)
 XX DE Artificial alpha-synuclein mutant A53T.

XX KW Artificial; alpha-synuclein mutant; accelerated aggregation;
 XX KW animal model; inhibitor; Parkinson's Disease.
 XX OS Homo sapiens.
 XX OS Synthetic.

XX PN W0200018917-A2.
 XX PD 06-APR-2000.

XX PF 27-SEP-1999; 99WO-US22302.
 XX PR 25-SEP-1998; 98US-0101862.
 XX PR 24-SEP-1999; 99US-0405035.

XX PA (AMGE-) AMGEN INC.
 XX PI Biere AL, Citron M;

XX WPI; 2000-293156/25.
 XX PT New artificial mutant alpha-synuclein polypeptides are useful to
 XX PT generate an animal model for Parkinson's disease and aggregate faster
 XX PT than the naturally occurring alpha-synuclein and alpha-synuclein
 XX PT mutants

XX PS Claim 3; Page 33; 34pp; English.
 XX CC AAY92090-92 are artificial alpha-synuclein mutant polypeptides which
 XX CC demonstrate accelerated aggregation compared to naturally occurring
 XX CC alpha-synuclein mutant polypeptides. Aggregate formation was enhanced
 XX CC for the E83Q-A90V and H50Y-A53T mutants, even when compared to the A53T
 XX CC mutant. The aggregate formation was significantly enhanced for the
 XX CC H50T-A53T-A76T mutant. An in vitro aggregation assay for evaluating
 XX CC alpha-synuclein mutants which demonstrate accelerated aggregation is
 XX CC also claimed. Additionally, a transgenic non-human mammal which has
 XX CC been genetically engineered to express the mutant alpha-synuclein
 XX CC polypeptide in the brain and an in vitro aggregation assay for high
 XX CC throughput screening of alpha-synuclein nucleation inhibitors are
 XX CC claimed. Alpha-synuclein inhibitors are used to treat, prevent or
 XX CC ameliorate Parkinson's Disease or related disorders (claimed). The
 XX CC transgenic animal is used as an animal model for Parkinson's Disease.

SQ Sequence 140 AA;


```
XX SQ Sequence 140 AA;
Query Match 97.4%; Score 684; DB 21; Length 140;
Best Local Similarity 97.9%; Pred. No. 2.2e-58;
Matches 137; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120

QY 121 DNEAYEMPSEEGYDYPEA 140
DB 121 DNEAYEMPSEEGYDYPEA 140

RESULT 14
AAY07274
ID AAY07274 standard; Protein; 140 AA.
AC AAY07274;
XX
XX
XX 06-JUL-1999 (first entry)
XX
XX Rat synuclein alternative gene splicing product #1.
XX Human; mouse; synuclein; persyn; diagnosis; neurodegenerative disorder;
XX cancer; breast; skin; intermediate filament damage.
XX
XX Rattus sp.
XX
XX EP908727-A1.
XX
XX 14-APR-1999.
XX
XX 21-SEP-1998; 98EP-0307628.
XX
XX 19-SEP-1997; 97GB-0019879.
XX
XX (NEUR-) NEUROPA LTD.
XX (UYSA-) UNIV ST ANDREWS.
XX
XX WPI; 1999-217169/19.
XX
XX New synuclein protein (persyn) and gene, useful in assays for
XX screening, diagnosing or monitoring cancer, neurodegenerative
XX disorders or skin disorders
XX
XX Disclosure; Page 20-21; 39pp; English.
XX
XX This sequence represents the sequence of a rat synuclein derived by
XX alternative gene splicing. The invention relates to novel human and
XX mouse synuclein family members designated persyn (AAY07271 and
XX AAY07172). The persyn sequence is useful for screening, diagnosing or
XX monitoring cancer (especially breast or skin cancer), neurodegenerative
XX disorders or skin disorders and for identifying cells having intermediate
XX filament damage.
XX
XX SQ Sequence 140 AA;
Query Match 95.4%; Score 670; DB 20; Length 140;
Best Local Similarity 95.0%; Pred. No. 4.8e-57;
Matches 133; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
```

```
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQMGKGEYPOEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYDYPEA 140
DB 121 SSEAYEMPSEEGYDYPEA 140

RESULT 15
ID AAY92220
XX AAY92220 standard; Protein; 140 AA.
AC AAY92220;
XX
XX 10-AUG-2000 (first entry)
XX
XX NACP/alpha-synuclein.
XX
XX NACP/alpha-synuclein; beta-synuclein; beta-synuclein delta 1; mutant;
XX primer; neurodegenerative; Alzheimers disease; Parkinsons disease;
XX Lewy body disease; transgene.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX Domain 73..83
XX FT /note= "missing from beta-synuclein"
XX
XX WO200020020-A2.
XX
XX 13-APR-2000.
XX
XX 06-OCT-1999; 99WO-US23134.
XX
XX 06-OCT-1998; 98US-0103310.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Masliah E;
XX
XX WPI; 2000-303637/26.
XX
XX Screening treatments for neurodegenerative disease, such as Alzheimer's
XX disease, comprises stimulating oxidative stress and measuring and
XX aggregation of non-amyloid beta component /alpha-synuclein in neuronal
XX cells
XX
XX Example 4; Page 47; 47pp; English.
XX
XX The invention concerns methods for screening treatments for, and
XX treatment of, neurodegenerative disease. The methods comprise measuring
XX aggregation of neurons of NACP/alpha-synuclein and stimulating expression
XX of a non-amyloidogenic protein in order to reduce the level of
XX aggregation. In the screening method, oxidative stress is stimulated in
XX the neuronal cells by introduction of metal ions and hydrogen peroxide.
XX The methods are useful for treating and screening treatments for
XX neurodegenerative disease consisting of Alzheimer's disease, Parkinson's
XX disease and Lewy body disease.
XX
XX SQ Sequence 140 AA;
Query Match 94.4%; Score 663; DB 21; Length 140;
Best Local Similarity 94.3%; Pred. No. 2.3e-56;
Matches 132; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60
DB 1 MDVFMKGLSKAKGEGVAAAEKTKGVAAAGKTKGVLVYVGSKTKEGVVHGVTVAEKT 60

QY 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQLGKNEGAPQEGILEDMPVDP 120
DB 61 EQVTNNGGAVVTGVTAVAQKTVEGAGTAATGFKVKKDQMGKGEYPOEGILEDMPVDP 120
```


OY 121 DNEAYEMPSEEGYQDYPEA 140
:|||||
Db 121 GSEAYEMPSEEGYQDYPEA 140

Search completed: January 28, 2003, 14:13:58
Job time : 8.80115 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 : Search time 1.83556 Seconds
(without alignments)
159.035 Million cell updates/sec

Title: US-09-904-987-4
Perfect score: 702
Sequence: 1 MDVFMKGLSKAKGEGVVAEAE.....DNEAYEMPSEGYQDYEPFA 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	100.0	140	US-09-904-987-4	Sequence 4, Appli
2	702	100.0	140	US-10-039-413-1	Sequence 1, Appli
3	695	99.0	140	US-10-039-413-2	Sequence 2, Appli
4	692	98.6	140	US-10-039-413-3	Sequence 3, Appli
5	684	97.4	140	US-10-039-413-4	Sequence 4, Appli
6	226	32.2	47	US-09-864-761-45900	Sequence 45900, A
7	151	21.5	28	US-09-864-761-46299	Sequence 46299, A
8	143	20.4	30	US-09-864-761-47813	Sequence 47813, A
9	83	11.8	321	US-09-967-347-12	Sequence 12, Appli
10	80.5	11.5	216	US-08-450-842-8	Sequence 8, Appli
11	79.5	11.3	129	US-09-815-242-11228	Sequence 11228, A
12	79	11.3	599	US-09-789-561-120	Sequence 120, App
13	78	11.1	267	US-09-852-555-2	Sequence 2, Appli
14	77	11.0	561	US-09-813-079A-2	Sequence 2, Appli
15	76	10.8	328	US-09-765-272-8	Sequence 8, Appli
16	76	10.8	475	US-09-734-569-166	Sequence 166, App
17	76	10.8	1599	US-10-092-880-9	Sequence 9, Appli
18	75.5	10.8	202	US-09-815-242-5604	Sequence 5604, Ap
19	75.5	10.8	203	US-09-815-242-12176	Sequence 12176, A

20	75	10.7	331	9	US-09-573-446-1	Sequence 1, Appli
21	75	10.7	487	10	US-09-745-763-108	Sequence 108, App
22	74.5	10.6	413	9	US-09-738-626-4478	Sequence 4478, Ap
23	73.5	10.5	714	10	US-09-978-242-3	Sequence 3, Appli
24	73	10.4	1781	10	US-09-738-877-3	Sequence 3, Appli
25	72.5	10.3	878	10	US-09-815-242-13957	Sequence 13957, A
26	72	10.3	401	10	US-09-815-242-13828	Sequence 13828, A
27	72	10.3	568	10	US-09-815-242-5140	Sequence 5140, Ap
28	72	10.3	836	9	US-09-738-626-5805	Sequence 5805, Ap
29	72	10.3	1536	9	US-10-092-880-2	Sequence 2, Appli
30	71.5	10.2	709	10	US-09-815-242-10989	Sequence 10989, A
31	71	10.1	811	10	US-09-840-787-45	Sequence 45, Appli
32	71	10.1	1002	9	US-09-988-117-3	Sequence 3, Appli
33	71	10.1	1002	10	US-09-812-471-3	Sequence 3, Appli
34	71	10.1	1002	10	US-09-812-633-3	Sequence 3, Appli
35	71	10.1	1477	9	US-10-092-880-4	Sequence 4, Appli
36	70.5	10.0	149	10	US-09-815-242-13554	Sequence 13554, A
37	70	10.0	338	10	US-09-815-242-13457	Sequence 13457, A
38	70	10.0	338	10	US-09-815-242-13645	Sequence 13645, A
39	70	10.0	486	10	US-09-825-414-7	Sequence 7, Appli
40	70	10.0	978	10	US-09-815-242-5456	Sequence 5456, Ap
41	70	10.0	1001	10	US-09-815-242-12686	Sequence 12686, A
42	70	10.0	1600	9	US-10-092-880-10	Sequence 10, Appli
43	70	10.0	2478	10	US-09-815-242-5816	Sequence 5816, Ap
44	70	10.0	2478	10	US-09-815-242-12967	Sequence 12967, A
45	69.5	9.9	484	10	US-09-815-242-10383	Sequence 10383, A

ALIGNMENTS

RESULT 1

US-09-904-987-4
; Sequence 4, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908Alactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep
; FILE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 140
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / XM_003494
; DATABASE ENTRY DATE: 2001-04-16
; RELEVANT RESIDUES: (1)..(140)
US-09-904-987-4

Query Match	100.0%	Score	702	DB	10	Length	140
Best Local Similarity	100.0%	Pred. No.	3.2e-58				
Matches	140	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MDVFMKGLSKAKGEGVVAEAEKTKQGVAAAGTKGKGVVGVGKTKGVVGVGVAETKTK 60					
Db	1	MDVFMKGLSKAKGEGVVAEAEKTKQGVAAAGTKGKGVVGVGKTKGVVGVGVAETKTK 60					
Qy	61	EQVTNVGAVTGVTVAAQKTVGAGSAAATGFKVKDQGLKNEGAPQEGILEMPVDP 120					
Db	61	EQVTNVGAVTGVTVAAQKTVGAGSAAATGFKVKDQGLKNEGAPQEGILEMPVDP 120					
Qy	121	DNEAYEMPSEGYQDYEPFA 140					
Db	121	DNEAYEMPSEGYQDYEPFA 140					
RESULT 2							
US-10-039-413-1							


```
; Sequence 1, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 140
; TYPE: PRF
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-1
Query Match 100.0%; Score 702; DB 12; Length 140;
Best Local Similarity 100.0%; Pred. No. 3.2e-58;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 3
US-10-039-413-2
; Sequence 2, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 140
; TYPE: PRF
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-2
Query Match 99.0%; Score 695; DB 12; Length 140;
Best Local Similarity 98.6%; Pred. No. 1.4e-57;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 4
US-10-039-413-3
; Sequence 3, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 140
; TYPE: PRF
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-3
Query Match 98.6%; Score 692; DB 12; Length 140;
Best Local Similarity 98.6%; Pred. No. 2.7e-57;
Matches 138; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
Db 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
QY 121 DNEAYEMPSEEGYQDYEP 140
Db 121 DNEAYEMPSEEGYQDYEP 140

RESULT 5
US-10-039-413-4
; Sequence 4, Application US/10039413
; Patent No. US20020152480A1
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; APPLICANT: Citron, Martin
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/10/039,413
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 60/101,862
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 140
; TYPE: PRF
; ORGANISM: ADULT HUMAN BRAIN
US-10-039-413-4
Query Match 97.4%; Score 684; DB 12; Length 140;
Best Local Similarity 97.9%; Pred. No. 1.5e-56;
Matches 137; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
Db 1 MDVFMKGLSKAKGVAAAEKTKQGVAAEAGTKEGVLYVGSKTKEGVHGVATVAEKT 60
QY 61 EQVTNVGGAVVTGTTVAQAQTVGAGSIAAATGFKKDLGKNEGAPQEGILEDMPVDP 120
```


Db 61 EQVTNNGGAVTGTAVTAAKTVGAGSIAAATGCFVKKDLGRNEGAPQEGILEMDPDP 120
Qy 121 DNEAYEMPSEEGYQDYPEA 140
Db 121 DNEAYEMPSEEGYQDYPEA 140

RESULT 6

US-09-864-761-45900
; Sequence 45900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45900
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015529.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: H70324.1, EVALUE 1.00e-08
; OTHER INFORMATION: SWISSPROT HIT: P37840, EVALUE 9.00e-10
US-09-864-761-45900

Query Match 32.2%; Score 226; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 7.6e-15;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 56 AEKTEQVTNNGGAVTGTAVTAAKTVGAGSIAAATGCFVKKDLGRNEGAPQEGILEMDPDP 102
Db 1 AEKTEQVTNNGGAVTGTAVTAAKTVGAGSIAAATGCFVKKDLGRNEGAPQEGILEMDPDP 47

RESULT 7

US-09-864-761-46299
; Sequence 46299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46299
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011025.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
; OTHER INFORMATION: EST_HUMAN HIT: AW793306.1, EVALUE 8.00e-10
; OTHER INFORMATION: SWISSPROT HIT: P37840, EVALUE 6.00e-11
US-09-864-761-46299

Query Match 21.5%; Score 151; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 NEEGAQEGILEDMVPDPNEAYEMPSE 130
|||||
Db 1 NEEGAQEGILEDMVPDPNEAYEMPSE 28

RESULT 8
US-09-864-761-47813
; Sequence 47813, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE REFERENCE: Aomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47813
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015529.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.56
; OTHER INFORMATION: SWISSPROT HIT: O55042, EVALUE 5.00e-10
; OTHER INFORMATION: EST_HUMAN HIT: BE567371.1, EVALUE 7.00e-09

US-09-864-761-47813

Query Match 20.4%; Score 143; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AKEGVVAAAEKTKQGVAAEAGTKKEGVLYV 40
|||||
Db 1 AKEGVVAAAEKTKQGVAAEAGTKKEGVLYV 30

RESULT 9
US-09-967-347-12
; Sequence 12, Application US/09967347
; Patent No. US20020115829A1
; GENERAL INFORMATION:
; APPLICANT: Finlay, B. Brett
; APPLICANT: Kenny, Brendan
; APPLICANT: Stein, Markus
; APPLICANT: Donnenberg, Michael S.
; APPLICANT: Lai, Li-Ching
; TITLE OF INVENTION: PATHOGENIC ESCHERICHIA COLI ASSOCIATED PROTEIN EspA
; FILE REFERENCE: 482112.401D1
; CURRENT APPLICATION NUMBER: US/09/967,347
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 09/171,517
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: PCT/CA97/00265
; PRIOR FILING DATE: 1997-04-23
; PRIOR APPLICATION NUMBER: 60/015,999
; PRIOR FILING DATE: 1996-04-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-967-347-12

Query Match 11.8%; Score 83; DB 10; Length 321;
Best Local Similarity 26.1%; Pred. No. 1.3;
Matches 36; Conservative 20; Mismatches 60; Indels 22; Gaps 6;

QY 6 KGLSKAKEGVAAAEKTKQGVAAEAGTKKEGV---LYVGSKTKEGVVHGVATVAEKTKEQ 62
|||
Db 150 KALAKASEGIAADAD-----DAAGAQQTATATAAKAASRT-SGITDDVATSAQKASQ- 200

QY 63 VTNVGGAVVTGTVTAQAQT-----VEGAGSTAATGFKDKDQKNEGAPQEGILEDM 116
|||
Db 201 ---VAEEAADAQAQAKAGLLSRFMAAAGRISGTPFIVVTSLAEGTKTLPTT-ISESV 256

QY 117 PVPDPNEAYEMPSEGYQ 134
|||
Db 257 KSNHDINEQRAKSVENLQ 274

RESULT 10
US-08-450-842-8
; Sequence 8, Application US/08450842
; Patent No. US20020045576A1
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:


```
RESULT 13
US-09-852-555-2
; Sequence 2, Application US/09852555
; Patent No. US20020034751A1
; GENERAL INFORMATION:
; APPLICANT: Yuen, kwok-Yung
; APPLICANT: Cao, Liang
; TITLE OF INVENTION: Aspergillus Fumigatus Antigenic Protein 1
; FILE REFERENCE: 609920-600015
; CURRENT APPLICATION NUMBER: US/09/852,555
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 60/203,322
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Aspergillus Fumigatus
US-09-852-555-2
Query Match 11.1%; Score 78; DB 10; Length 267;
Best Local Similarity 26.4%; Pred. No. 3;
Matches 24; Conservative 20; Mismatches 41; Indels 6; Gaps 3;
QY 5 MKGLSKAKGVV-AAAEKTKQGVAAEAGTKGVLVVGSKTKGVVHGVA-TVAEKTKEQ 62
Db 76 VQDTUKQVEGVDDLLSKDKKFFVANNAG----GTVYEDLKAQYTAADSLAKAISAKVPES 131
QY 63 VTNVGGAVVTGVTAVAAQKTVEGAGSTAAATG 93
Db 132 LSDIAAQLSAGITAAIQKIDAYKDAASSTG 162
RESULT 14
US-09-812-079A-2
; Sequence 2, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-079A-2
Query Match 11.0%; Score 77; DB 10; Length 561;
Best Local Similarity 26.5%; Pred. No. 9.2;
Matches 26; Conservative 17; Mismatches 45; Indels 10; Gaps 3;
QY 28 EAAKTKGVLVVGSKTKGVVHGVAETKKEQVTNNGAVTGTAVAAQKTVEGAGS 87
Db 390 EGAAGDDDDVTSGSDSEELV-----TTERKTPRVTG-GGAMAGASTSAGRKRKSASSA 443
QY 88 IAAATGFVKKDKLCKNEGAPQEGILEMDPVPDPNEAY 125
Db 444 TACTSGVMTGRGLKRAESTVAPE-----EDTDEDSONEIH 477
RESULT 15
US-09-765-272-8
; Sequence 8, Application US/09765272
```

```
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-765-272-8
Query Match 10.8%; Score 76; DB 10; Length 328;
Best Local Similarity 30.8%; Pred. No. 5.9;
Matches 36; Conservative 12; Mismatches 55; Indels 14; Gaps 5;
QY 16 VAAAEKTKQGVAAEAGTKGVLVVGSKTKGVVHGVAETKKEQVTNNG--GAVVTG 73
Db 132 VAAAKTTTKQGVFVGGIESEVI---SREFAGFKAGVASVDFSIKQVDYAGSFGDAAKG 188
QY 74 VTAVAQKTVEGAG---STAAATG---FVKKDKLCKNEGAPQEGILEMDPVPDPNEA 124
Db 189 KTIATAAQYAAGADIVYQVAGGTGAGVFAEAKSL---NESRPNENKVVWVIGVDRDQEA 242
Search completed: January 28, 2003, 14:40:08
Job time : 2.83556 secs
```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 2.90631 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-4

Perfect score: 702

Sequence: 1 MDVFMKGLSKAKEGVVAAAE.....DNEAYEMPSEGYQDYEPFA 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep:*
- 4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep:*
- 5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	702	100.0	140	US-09-405-035-1	Sequence 1, Appli
2	702	100.0	140	PCT-US94-09789-2	Sequence 2, Appli
3	695	99.0	140	US-09-405-035-2	Sequence 2, Appli
4	692	98.6	140	US-09-405-035-3	Sequence 3, Appli
5	694	97.4	140	US-09-405-035-4	Sequence 4, Appli
6	307.5	43.8	127	US-08-705-771-12	Sequence 12, Appli
7	166	23.6	35	PCT-US94-09789-3	Sequence 3, Appli
8	95	13.5	20	PCT-US94-09789-4	Sequence 4, Appli
9	93	13.2	216	US-09-125-619-46	Sequence 46, Appli
10	88	12.5	214	US-09-125-619-34	Sequence 34, Appli
11	87	12.4	190	US-09-125-619-15	Sequence 15, Appli
12	87	12.4	212	US-09-125-619-31	Sequence 31, Appli
13	87	12.4	212	US-09-125-619-43	Sequence 43, Appli
14	87	12.4	356	US-09-125-619-2	Sequence 2, Appli
15	87	12.4	356	US-09-125-619-13	Sequence 13, Appli
16	85.5	12.2	189	US-09-125-619-26	Sequence 26, Appli
17	83	11.8	321	US-09-171-5178-12	Sequence 12, Appli
18	82.5	11.8	187	US-09-125-619-23	Sequence 23, Appli
19	80.5	11.5	190	US-09-125-619-19	Sequence 19, Appli
20	80.5	11.5	216	US-08-451-947-8	Sequence 8, Appli
21	80.5	11.5	216	US-08-424-826A-8	Sequence 8, Appli
22	80.5	11.5	216	US-08-928-694-8	Sequence 8, Appli
23	80.5	11.5	216	PCT-US91-06950-8	Sequence 8, Appli
24	79.5	11.3	211	US-09-125-619-40	Sequence 40, Appli
25	79.5	11.3	212	US-09-125-619-41	Sequence 41, Appli
26	79	11.3	194	US-09-125-619-28	Sequence 28, Appli
27	78.5	11.2	281	US-08-405-175A-9	Sequence 9, Appli

28	78	11.1	212	4	US-09-125-619-32	Sequence 32, Appli
29	77.5	11.0	179	4	US-09-125-619-18	Sequence 18, Appli
30	77.5	11.0	209	4	US-09-125-619-42	Sequence 42, Appli
31	77.5	11.0	729	4	US-08-887-534A-30	Sequence 30, Appli
32	77.5	11.0	1004	4	US-09-268-347-30	Sequence 30, Appli
33	77	11.0	579	4	US-09-171-699-8	Sequence 8, Appli
34	76.5	10.9	215	4	US-09-125-619-36	Sequence 36, Appli
35	76.5	10.9	892	4	US-09-336-447A-5	Sequence 5, Appli
36	76.5	10.9	1477	4	US-09-206-942-71	Sequence 71, Appli
37	76	10.8	158	4	US-09-125-619-21	Sequence 21, Appli
38	76	10.8	195	4	US-09-125-619-17	Sequence 17, Appli
39	76	10.8	209	4	US-09-125-619-33	Sequence 33, Appli
40	76	10.8	211	4	US-09-125-619-37	Sequence 37, Appli
41	76	10.8	213	4	US-09-125-619-39	Sequence 39, Appli
42	76	10.8	328	4	US-08-961-083-8	Sequence 8, Appli
43	76	10.8	579	4	US-09-171-699-6	Sequence 6, Appli
44	76	10.8	1338	2	US-08-728-470-9	Sequence 9, Appli
45	76	10.8	1338	4	US-08-719-641-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-405-035-1
; Sequence 1, Application US/09405035
; Patent No. 6184351
; GENERAL INFORMATION:
; APPLICANT: Biere, Anja Leona
; TITLE OF INVENTION: ALPHA-SYNUCLEIN SUPER-MUTANTS ACCELERATE
; TITLE OF INVENTION: ALPHA-SYNUCLEIN AGGREGATION
; FILE REFERENCE: A-565
; CURRENT APPLICATION NUMBER: US/09/405.035
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 60/101,862
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1
; LENGTH: 140
; TYPE: PRT
; ORGANISM: ADULT HUMAN BRAIN
US-09-405-035-1

Query Match		100.0%;	Score 702;	DB 4;	Length 140;
Best Local Similarity		100.0%;	Pred. No. 3.9e-68;		
Matches 140;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MDVFMKGLSKAKEGVVAAAEKTKQGVAAAGTKEGVLYVGSKTKEGVVHGVATVAEKT 60			
Db	1	MDVFMKGLSKAKEGVVAAAEKTKQGVAAAGTKEGVLYVGSKTKEGVVHGVATVAEKT 60			
Qy	61	EQVTNVGGAVTGTAVTAQKTEGAGSIAAATGFVKKDQLGKNEGAPQEGILEMPVDP 120			
Db	61	EQVTNVGGAVTGTAVTAQKTEGAGSIAAATGFVKKDQLGKNEGAPQEGILEMPVDP 120			
Qy	121	DNEAYEMPSEGYQDYEPFA 140			
Db	121	DNEAYEMPSEGYQDYEPFA 140			

RESULT 2

PCT-US94-09789-2
; Sequence 2, Application PC/TUS9409789
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN
; TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500

LENGTH: 140

TYPE: PRT

ORGANISM: ADULT HUMAN BRAIN

US-09-405-035-4

Query Match 97.4%; Score 684; DB 4; Length 140;
Best Local Similarity 97.9%; Pred. No. 3.3e-66;
Matches 137; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVFMKLSKAKGVAAGAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVVGHVATVAEKT 60
|||||

Db 1 MDVFMKLSKAKGVAAGAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVVGHVATVAEKT 60
|||||

QY 61 EQVTNVGGAVVTGVTVAQAQKTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120
|||||

Db 61 EQVTNVGGAVVTGVTVAQAQKTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120
|||||

QY 121 DNEAYEMPSEEGYQDYPEA 140
|||||

Db 121 DNEAYEMPSEEGYQDYPEA 140
|||||

RESULT 6

US-08-705-771-12

Sequence 12, Application US/08705771

Patent No. 6054289

GENERAL INFORMATION:

APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,

APPLICANT: Jian Ni and Jing-Shan Hu

TITLE OF INVENTION: Human Genes, Sequences and

TITLE OF INVENTION: Expression Products

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/705,771

FILING DATE: August 30, 1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-346 (PF196)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 973-994-1700

TELEFAX: 973-994-1744

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 127 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-705-771-12

Query Match 43.8%; Score 307.5; DB 3; Length 127;
Best Local Similarity 61.8%; Pred. No. 8.5e-26;
Matches 68; Conservative 10; Mismatches 29; Indels 3; Gaps 1;

QY 1 MDVFMKLSKAKGVAAGAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVVGHVATVAEKT 60
|||||

Db 1 MDVFMKLSKAKGVAAGAAEKTKQGVAAAGKTKGVLVYVGSKTKEGVVGHVATVAEKT 60
|||||

QY 61 EQVTNVGGAVVTGVTVAQAQKTVGAGSIAAATGFVKKDLGKNEGAPQEGILEDMPVDP 120
|||||

Db 61 EQANAVSKAVSSVNTVATKTVEEAENIAVTSVGVVRKEDL---RPSAPOO 107
|||||

RESULT 7

PCT-US94-09789-3

Sequence 3, Application PC/TUS9409789

GENERAL INFORMATION:

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN

TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09789

FILING DATE: 29-AUG-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Howells, Stacy L.

REGISTRATION NUMBER: 34,842

REFERENCE/DOCKET NUMBER: FD-3520

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

CLONE: NAC

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..35

PCT-US94-09789-3

Query Match

Best Local Similarity 23.6%; Score 166; DB 5; Length 35;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EQVTNVGGAVVTGVTVAQAQKTVGAGSIAAATGFV 95
|||||

Db 1 EQVTNVGGAVVTGVTVAQAQKTVGAGSIAAATGFV 35
|||||

RESULT 8

PCT-US94-09789-4

Sequence 4, Application PC/TUS9409789

GENERAL INFORMATION:

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: NOVEL COMPONENT OF AMYLOID IN

TITLE OF INVENTION: ALZHEIMER'S DISEASE AND METHODS FOR USE OF SAME

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz

STREET: 1880 Century Park East - Suite 500

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90067

Query Match	12.4%	Score 87;	DB 4;	Length 212;
Best Local Similarity	33.3%	Pred. No. 0.095;		
Matches 38;	Conservative 14;	Mismatches 44;	Indels 18;	Gaps 7;
QY	5	MKGLSKAKGVVAAA----	EKTQKGVAEAAKTKGVLVYVSGTKGVVHGVATVAEKTKE	61

Query Match	12.4%	Score 87;	DB 4;	Length 356;
Best Local Similarity	33.3%;	pred. No. 0.17;		
Matches	38;	Conservative 14;	Mismatches 44;	Indels 18; Gaps 7;
Qy	5	MKGLSKAKEGVVAAA---EKTQGVAAAGKTKGCVLVGSKTKRGVVHGVGVATVAEKTKE	61	

Db 164 VKGIAGIKIETVEAAGGSEKL-AVAAAKGENNKGAGKLFKA-GAAAHGDSEASKAAG 221
QY 62 QVTNWGG-----AVVTGVTVAQ---KTVEGAGSIAAATGFVKKD---OLGKNE 104
Db 222 AVSAVSGEQILSAIVTAADAAEQDGKKPEAKNPAAAAIG--DKOGGAEEFGODE 273

Search completed: January 28, 2003, 14:21:22
Job time : 3.90631 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 19.8765 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPKQEFVEMDHAGTYGL.....SPQLATLADVSASLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-73.*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2780.5	70.1	733	2 A45301	microtubule-associ
2	2552	64.4	686	2 A38235	microtubule-associ
3	2116.5	53.4	441	1 QRH071	microtubule-associ
4	1845	46.5	432	2 J50306	microtubule-associ
5	1845	46.5	448	1 ORB071	microtubule-associ
6	1770	44.6	402	1 ORB072	microtubule-associ
7	1552	39.1	374	2 S46264	microtubule-associ
8	1415.5	35.7	316	1 ORH072	microtubule-associ
9	1354.5	34.2	341	2 B28820	microtubule-associ
10	1350.5	34.1	364	2 A28820	microtubule-associ
11	812.5	20.5	1824	1 ORH071	microtubule-associ
12	807	20.4	1828	2 A40115	microtubule-associ
13	796.5	20.1	1830	2 A37981	microtubule-associ
14	795	20.1	1825	2 S13507	microtubule-associ
15	752	19.0	472	2 I67793	microtubule-associ
16	715	18.0	323	2 A55983	microtubule-associ
17	624	15.7	1152	2 A33183	microtubule-associ
18	621.5	15.7	1125	2 B41206	microtubule-associ
19	576	14.5	381	2 S51375	microtubule-associ
20	545	13.7	1072	2 A37127	microtubule-associ
21	507.5	12.8	982	2 A53253	microtubule-associ
22	482.5	12.2	1224	2 T14007	microtubule-associ
23	481	12.1	125	2 I52650	microtubule-associ
24	341	8.6	66	2 S26663	microtubule-associ
25	309	7.8	2187	2 T30826	nascent polypeptid
26	294	7.4	1464	2 S59856	collagen alpha 1(I
27	290	7.3	1049	1 CGB07S	collagen alpha 1(I
28	285.5	7.2	990	2 I51618	nucleolar phosphop
29	285.5	7.2	1806	1 CGHU1E	collagen alpha 1(X

30	285	7.2	971	2 T19431	hypothetical prote
31	285	7.2	1184	2 G01763	atrophin-1 - human
32	283.5	7.2	779	1 CGB01S	collagen alpha 1(I
33	283.5	7.2	1466	1 CGHU7L	collagen alpha 1(I
34	283	7.1	1838	1 CGHU1V	collagen alpha 1(V
35	282	7.1	1151	2 T18535	high molecular mas
36	281.5	7.1	886	2 I50894	collagen alpha 1(I
37	279.5	7.0	1691	1 S22917	collagen alpha 5(I
38	278	7.0	242	2 T16349	hypothetical prote
39	277.5	7.0	1763	2 S16366	hypothetical prote
40	277.5	7.0	2715	2 T13049	eyelid - fruit fly
41	276	7.0	880	2 D89756	protein T23E7.2b1
42	276	7.0	1274	2 T16251	hypothetical prote
43	276	7.0	1492	2 A40333	collagen alpha 1(I
44	275	6.9	1024	2 S18251	collagen alpha 1(X
45	275	6.9	1549	2 I48103	type VII collagen

ALIGNMENTS

RESULT 1

A45301

microtubule-associated protein tau - mouse

N:Alternate names: microtubule binding protein tau

C:Species: Mus musculus (house mouse)

C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 13-Aug-1999

C:Accession: A45301; S31658

R:Couchie, D.; Mavilia, C.; Georgieff, I.S.; Liem, R.K.; Shelanski, M.L.; Nunez, J.

A:Title: Primary structure of high molecular weight tau present in the peripheral ner

A:Reference number: A45301; MUID:92262443; PMID:1374898

A:Accession: A45301

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-733 <COU>

A>Note: this sequence is inconsistent with the nucleotide translation

A>Note: sequence extracted from NCBI backbone (NCBI:102045, NCBI:102046)

R:Kenner, L.; Forstner, M.; Hutter, H.; Hoefler, G.; Kurzbaue, R.; Zatloukal, K.; Kr

A:Description: First observation of mRNA for a tau-protein from murine liver and kidn

A:Reference number: S31658

A:Accession: S31658

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-529-651 <KEN>

A:Cross-references: EMBL:Z12133; NID:g54263; PIDN:CAA78121.1; PID:g388534

C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology

C:Keywords: microtubule binding; tandem repeat

F:544-574/Domain: MAP2/tau repeat homology <MT1>

F:575-605/Domain: MAP2/tau repeat homology <MT2>

F:606-636/Domain: MAP2/tau repeat homology <MT3>

F:637-668/Domain: MAP2/tau repeat homology <MT4>

Query Match

Best Local Similarity 70.1%; Score 2780.5; DB 2; Length 733;

Matches 569; Conservative 44; Mismatches 111; Indels 43; Gaps 11;

QY 1

MAEPKQEFVEMDHAGTYGLDRKDGQYTMHODQEGDTDAGLKESPLQTPPTDGSSEEPG 60

Db 1

MANPROEFTMEDHA-----GDTLLDQEGDMHGLKESPPQPPADGAEPEG 49

QY 61

SETSDAKSTTAEVDVAPLVDEGAPKQAAAPHTTEIGETTAEEAGIGTPTPLEDEAAG 120

Db 50

SETSDAKSTTAEVDVAPLVDERAPKQAAAPHTTEIGETTAEEAGIGTPTPLEDEAAG 109

QY 121

HVTPQEPGKVGVEGFLREPGLSHOLMSGMPGLLPEGPREATROPSPGTCPEDETEG 180

Db 110

HVTPQ-----GR--REGQADPLGTSDWTRQQVSSMSGAPLLPQGLREATCQPSGTRPEDIEK 163

QY 181

GRHAPELLKQLGLDHLHQEGPPLKGGAGKRRPGSKKEVEDRDVDDESSQDSSPPSKASPA 240

Db 164

SHPASELLR-----RGPPQKEGWDRLGSEEEVDEDLTVDESS-QDSPPSQASLT 213

A:Accession: JN0009
A:Molecule type: mRNA
A:Residues: 1-44,103-274,306-441 <LPE>
R:Goedert, M.; Spillantini, M.G.; Potier, M.C.; Ulrich, J.; Crowther, R.A.
EMBO J. 8, 393-399, 1989
A:Title: Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated
A:Reference number: S03796; MUID:89251564; PMID:2498079
A:Accession: S03796
A:Molecule type: mRNA
A:Residues: 1-44,103-441 <CO>
A:Cross-references: EMBL:X14474; NID:g36724; PIDN:CAA32636.1; PID:g36725
R:Andreasson, A.; Brown, W.M.; Kosik, K.S.
Biochemistry 31, 10626-10633, 1992
A:Title: Structure and novel exons of the human tau gene.
A:Reference number: S26662; MUID:93041757; PMID:1420178
A:Accession: S26662
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 144-185 <AND>
A:Cross-references: EMBL:X61372; NID:g36718; PID:g36719
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
A:Accession: S26666
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 187-274 <AN2>
A:Cross-references: EMBL:X61374; NID:g36722; PID:g36723
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
A:Accession: S26662
A:Molecule type: DNA
A:Residues: 371-441 <ANW>
R:Jakes, R.; Novak, M.; Davison, M.; Wischik, C.M.
EMBO J. 10, 2725-2729, 1991
A:Title: Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's disease
A:Reference number: S17302; MUID:92007714; PMID:1915258
A:Accession: S17302
A:Status: preliminary
A:Molecule type: protein
A:Residues: 268-274,306-395 <JAK>
R:Hasegawa, M.; Morishima-Kawashima, M.; Takio, K.; Suzuki, K.; Ihara, Y.
J. Biol. Chem. 267, 17047-17054, 1992
A:Title: Protein sequence and mass spectrometric analyses of tau in the Alzheimer's disease
A:Reference number: A4344; MUID:92381012; PMID:1512244
A:Accession: A4344
A:Molecule type: protein
A:Residues: 2-73,103-130,151-180,191-254,260-269,275-290,299-317,322-340,344-347,354-383
A:Experimental source: Alzheimer's disease brain
A:Note: sequence extracted from NCBI backbone (NCBIP:112039)
C:Comment: This heterogeneous protein, which is found predominantly in cells of the nervous system, is the core protein of the paired helical filament of Alzheimer's disease.
C:Genetics:
A:Gene: GDB:MAPT
A:Cross-references: GDB:119434; OMIM:157140
A:Map position: 17q21-17q21
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; Alzheimer's disease; duplication; microtubule binding;
F:1-441/Product: microtubule-associated protein tau, long splice form #status predicted
F:1-274,306-441/Product: microtubule-associated protein tau (clone htau39) #status predicted
F:1-73,103-441/Product: microtubule-associated protein tau (clone htau34) #status predicted
F:1-73,103-274,306-441/Product: microtubule-associated protein tau (clone htau34) #status predicted
F:1-44,103-274,306-441/Product: microtubule-associated protein tau (clone htau37) #status predicted
F:1-44,103-441/Product: microtubule-associated protein tau type II #status predicted <MA
F:252-282/Domain: MAP2/tau repeat homology <WT1>
F:283-313/Domain: MAP2/tau repeat homology <WT2>
F:314-344/Domain: MAP2/tau repeat homology <WT3>
F:345-376/Domain: MAP2/tau repeat homology <WT4>
Query Match 53.4%; Score 2116.5; DB 1; Length 441;
Best Local Similarity 58.2%; Pred. No. 7,9e-90;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGRKQGGYTMHQDEGDTAGLKESPLQTPTEDESGEPG 60
|||||

Db 1 MAEPQEFVEMDHAGTYGLGRKQGGYTMHQDEGDTAGLKESPLQTPTEDESGEPG 60
QY 61 SETSDAKSTPTAEDVTALVDEGAPGKAAQAQPHTEIPEGTTAEAGTGTPTPSLEDEAAG 120
|||||
Db 61 SETSDAKSTPTAEDVTALVDEGAPGKAAQAQPHTEIPEGTTAEAGTGTPTPSLEDEAAG 120
QY 121 HVTQBPESGKVVQEGFLEPGPGGLSHQMGPCAPLLPEGPREATRQPSGTGPDTEG 180
|||||
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHQLLDLHOGEPPLKAGAGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIQCPAEGAIPLVDFLSKVSTIPIASEPDGSPVGRAGQODAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEQHSEHGLRAAPGAPGEGPEARGPSILGDTKEADLPEPSEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQPKARVMKSKDGTGSDDKAKKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
|||||
Db 125 -----ARVMKSKDGTGSDDKA----- 142
QY 421 DPLQPSPPAVCPPEPPSPKIVSVTSRTGSGAKEMKLGADGKTATPRGAAPGQK 480
|||||
Db 143 -----KGADGKTATPRGAAPGQK 163
QY 481 GOANATRIPTAKTPAPKTPPPSSGEPKSGDRSGYSSPGSGTSGSRRTPSLPTPTPTREP 540
|||||
Db 164 GOANATRIPTAKTPAPKTPPPSSGEPKSGDRSGYSSPGSGTSGSRRTPSLPTPTPTREP 223
QY 541 KKVAVVTRPPKSPSSAKSRLQTPVMPDLKNVSKIGSTENLKHQGGKVKQIINKKLD 600
|||||
Db 224 KKVAVVTRPPKSPSSAKSRLQTPVMPDLKNVSKIGSTENLKHQGGKVKQIINKKLD 283
QY 601 LSNVQSKCGSKDNKIHVPGGSGVQIVYKPVDSKVTSCGSLGNIHHKPGGQGVSEK 660
|||||
Db 284 LSNVQSKCGSKDNKIHVPGGSGVQIVYKPVDSKVTSCGSLGNIHHKPGGQGVSEK 343
QY 661 LDFKDRVQSKTGSIDNITHVPGGGNNKTEHKLTFRENAKAKTDHGAIVYKSPVWSDGT 720
|||||
Db 344 LDFKDRVQSKTGSIDNITHVPGGGNNKTEHKLTFRENAKAKTDHGAIVYKSPVWSDGT 403
QY 721 SPRHLNSVSTGSDMDVDSPLATLADSVASLAKQGL 758
|||||
Db 404 SPRHLNSVSTGSDMDVDSPLATLADSVASLAKQGL 441
RESULT 4
JS0306
microtubule-associated protein tau - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 31-Dec-1993
C:Accession: JS0306; A33574
R:Kosik, K.S.; Orcschio, L.D.; Bakalis, S.; Neve, R.L.
Neuron 2, 1389-1397, 1989
A:Title: Developmentally regulated expression of specific tau sequences.
A:Reference number: JS0306; MUID:90180457; PMID:2560640
A:Accession: JS0306
A:Molecule type: mRNA
A:Residues: 1-432 <KOS>
A:Note: the sequence shown is from adult rat brain
A:Note: the partial sequence from fetal rat brain is lacking residues 266-296; the fe
R:Kanai, Y.; Takemura, R.; Oshima, T.; Mori, H.; Ihara, Y.; Yanagisawa, M.; Masaki, T.
J. Cell Biol. 109, 1173-1184, 1989
A:Title: Expression of multiple tau isoforms and microtubule bundle formation in fibr
A:Reference number: A33574; MUID:89359509; PMID:2504728
A:Accession: A33574
A:Status: not compared with conceptual translation

QY 181 GRHAPPELLKHQLLDGLHQEGPLKAGGKERPGSKEEVEDRVDSESSPQSPKASPA 240
Db 114 -----
QY 241 QDGRPPOTAAREATSIPGPPAEGAIPLPVDFLSKVSVEIPASEPDPGSGVRAKQADAPLE 300
Db 114 -----
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPGAPGEGPEARGPSLIGEDTKADLPPESEKQPA 360
Db 114 -----
QY 361 AAPRGKPSRVPLKARMYSKSDGTGDDKAKTSTRSSAKTLKRNPCLSPKLPTPGSS 420
Db 114 -----
QY 421 DPLIQSSPAVCPPEPPSPKHVSSTRTGSSGAKEMKLGADGK--TKIATPRGAAPP 478
Db 132 -----
QY 479 QKGOANATRIAPAKTPAPKTPPSS-----GPPKSGDRSGSYSSPGSPGT 522
Db 153 QKGOANATRIAPAKTTPPTPKTSPATMQVKPPAGAKSERGESGSGDRSGSYSSPGSPGT 212
QY 523 PGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQTAPVPMPLKNVSKSIGSTEN 582
Db 213 PGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQAAPCPMDLKNVSKSIGSTEN 272
QY 583 LKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVDSLKVTSKCGSL 642
Db 273 LKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPVDSLKVTSKCGSL 332
QY 643 GNTHKPGGGQVEVSEKLDKDFKRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAK 702
Db 333 GNTHKPGGGQVEVSEKLDKDFKRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAK 392
QY 703 TDHGAIEIVYKSPVSGDTSRPHLSNVSTGSDIMVDSPLQATLADEVSAKQGL 758
Db 393 TDHGAIEIVYKSPVSGDTSRPHLSNVSTGSDIMVDSPLQATLADEVSAKQGL 448
RESULT 6
QRB02
N:Contains: microtubule-associated protein tau, form 3 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 31-Mar-1996
R:Himmeler, A.; Drechsel, D.; Kirschner, M.W.; Martin Jr., D.W.
Mol. Cell. Biol. 9, 1381-1388, 1989
A:Title: Tau consists of a set of proteins with repeated C-terminal microtubule-binding
A:Reference number: A31939; MUID:89261765; PMID:2498649
A:Accession: B31939
A:Molecule type: mRNA
A:Residues: 1-402 <HIM>
A:Cross-references: GB:M26157; GB:M26158
R:Paudel, H.K.; Lew, J.; Ali, Z.; Wang, J.H.
J. Biol. Chem. 268, 23512-23518, 1993
A:Title: Brain proline-directed protein kinase phosphorylates tau on sites that are abno
A:Reference number: A48885; MUID:94043150; PMID:8226879
A:Accession: A48885
A:Molecule type: protein
A:Residues: 'X', 157-162, 'X', 164-165, 'X', 167-170, 192-195, 'X', 197-201, 'X', 358-364, 'X', 366-
A:Experimental source: brain
A:Note: sequence modified after extraction from NCBI backbone
R:Aizawa, H.; Kawasaki, H.; Murofushi, H.; Kotani, S.; Suzuki, K.; Sakai, H.
J. Biol. Chem. 263, 7703-7707, 1988
A:Title: Microtubule-binding domain of Tau proteins.
A:Reference number: A28173; MUID:88227970; PMID:3131325
A:Accession: A28173
A:Molecule type: protein
A:Residues: 159-172, 'X', 174-177 <AIZ>

A:Experimental source: brain
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; phosphoprotein; tandem repeat
F:1-402/Product: microtubule-associated protein tau, form 3 #status predicted <Bf4>
F:1-243,297-402/Product: microtubule-associated protein tau, form 5 #status predicted
F:101-402/Product: microtubule-associated protein tau, form 4 #status predicted <Bf7>
F:159-177/Region: microtubule binding #status experimental
F:213-243/Domain: MAP2/tau repeat homology <WT1>
F:244-274/Domain: MAP2/tau repeat homology <WT2>
F:275-305/Domain: MAP2/tau repeat homology <WT3>
F:306-337/Domain: MAP2/tau repeat homology <WT4>
F:156,163,196,202,365/Binding site: phosphate (ser) (covalent) (by proline-directed k
F:166/Binding site: phosphate (Thr) (covalent) (by proline-directed kinase) #status p

Query Match 44.6%; Score 1770; DB 1; Length 402;
Best Local Similarity 94.6%; Pred. No. 4.5e-74;
Matches 347; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

QY 394 KTSTRSAKTLKRNPCLSPKLPTPGSSDPLIQSPSPAVCPPEPPSPKHVSSTRTGSSG 453
Db 38 KPSTPSSAKTLKRNPCLSPKRPTGSSDPLIKPSPAVCPPEPPSPKHVSSTRTGNSG 97
QY 454 AKEMKLGADGK--TKIATPRGAAPPQKGOANATRIAPAKTPPAPKTPPSSGEPKSGDR 511
Db 98 AKEMKVGADGKPGCTKIATPRGAAPPQKGOANATRIAPAKTTPPTKTP--GESGKSGDR 155
QY 512 SGYSSPGSPGTPGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQTAPVPMPLDK 571
Db 156 SGYSSPGSPGTPGSRRTPSLPTPTREPCKVAVRTPPKSPSAKSRLQAAPGMPDLK 215
QY 572 NVKSKIGSTENLKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPV 631
Db 216 NVKSKIGSTENLKHQPGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSGVQIVYKPV 275
QY 632 LSKVTSKCGSLGNTHHKPGGQVEVSEKLDKDFKRVQSKIGSLDNITHVPGGNGKKIETH 691
Db 276 LSKVTSKCGSLGNTHHKPGGQVEVSEKLDKDFKRVQSKIGSLDNITHVPGGNGKKIETH 335
QY 692 KLTFRENAKAKTDHGAIEIVYKSPVSGDTSRPHLSNVSTGSDIMVDSPLQATLADEVSA 751
Db 336 KLTFRENAKAKTDHGAIEIVYKSPVSGDTSRPHLSNVSTGSDIMVDSPLQATLADEVSA 395
QY 752 SLAKQGL 758
Db 396 SLAKQGL 402

RESULT 7
S46264
microtubule-associated protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 13-Aug-1999
R:Sadot, E.; Marx, R.; Barg, J.; Behar, L.; Ginzburg, I.
J. Mol. Biol. 241, 325-331, 1994
A:Title: Complete sequence of 3'-untranslated region of tau from rat central nervous
A:Reference number: S46264; MUID:94334997; PMID:8057376
A:Accession: S46264
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <SAD>
C:Cross-references: EMBL:X79321; NID:g517393; PIDN:CAA55889.1; PID:g517394
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
F:185-215/Domain: MAP2/tau repeat homology <WT1>
F:216-246/Domain: MAP2/tau repeat homology <WT2>
F:247-277/Domain: MAP2/tau repeat homology <WT3>
F:278-309/Domain: MAP2/tau repeat homology <WT4>

Query Match 39.1%; Score 1552; DB 2; Length 374;
Best Local Similarity 45.3%; Pred. No. 3.6e-64;
Matches 344; Conservative 14; Mismatches 14; Indels 388; Gaps 5;

QY 1 MAEPQREFEVMEHDHAGTYGLGDRKDQGGTYMHQDEGDTAGLKESPLQPTPTDSEEPG 60


```
|||||: ||| | | ||| |||| | |||
Db 1 MAEPQEFDTMEDQA-----GDYTLQDQEGMDHGLK----- 33
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAQPHTEIPGTTAREAGIGDTPSLEDAAG 120
Db 34 -----AEEAGIGDTPNMEDQAAG 51
QY 121 HVTQEPESKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDLHOEGPLKAGGERPGSKEEVEDRDVDESSPQSPSKASPA 240
Db 56 ----- 55
QY 241 QDGRPPQTAAREATSI PGFPAEGA IPLVDFLSKVSTEIPASEPDGPSVGRAGGADAPLE 300
Db 56 ----- 55
QY 301 FTFHVEITPNVQEQAHSEHILGRAAFPAGGEGPEARGPSLIGEDTKADLPSEKOPA 360
Db 56 ----- 55
QY 361 AAPRGKPSRVQOLKARMVSKSGDGTGDDKAKTSTRSSAKTLKNRCLSPKLPFGSS 420
Db 56 -----ARVAGSKDRTGDERKA----- 73
QY 421 DPLIQSSPAVCEPPSPKPKHSSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 74 -----KGADGKTGAKIATPRGAATPG 94
QY 479 QKQANATRIAPKTPPSSGPPKSGDRSGYSPGSPGTPGSRSTRPSLTPPTR 538
Db 95 QKGTSNATRIAPKTPPSPGSGEPKSGERSGYSPGSPGTPGSRSTRPSLTPPTR 154
QY 539 EPKKAIVVTPPKSPSSAKSRLOTPAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINK 598
Db 155 EPKKAIVVTPPKSPSSAKSRLOTPAPVMPDLKNVRSKIGSTENLKHQPGGKQVQIINK 214
QY 599 LQLSNVQSKGSKDNIKHVPGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVKS 658
Db 215 LQLSNVQSKGSKDNIKHVPGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVKS 274
QY 659 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVVS 718
Db 275 EKLDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVVS 334
QY 719 DTSPRHLSNVSTGSDMDVDSPOLATLADEVSAKQGL 758
Db 335 DTSPRHLSNVSTGSDMDVDSPOLATLADEVSAKQGL 374

RESULT 8
QRHUT2
microtubule-associated protein tau, fetal (clone p18) - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 02-Sep-1997
C:Accession: P00001
R:Lee, G.; Neve, R.L.; Kosik, K.S.
Neuron 2, 1615-1624, 1989
A:Title: The microtubule binding domain of tau protein.
A:Reference number: JN0009; MUID:90180482; PMID:2516729
A:Accession: P00001
A:Molecule type: mRNA
A:Residues: 1-316 <LEE>
A>Note: this sequence differs from a previously reported fetal tau protein sequence only
C:Genetics:
A:Gene: GDB:MAPT; MTBT1
A:Cross-References: GDB:119434; OMIM:157140
A:Map position: 17q21-17q21
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; Alzheimer's disease; duplication; microtubule binding;
F:158-188/Domain: MAP2/tau repeat homology <MT1>
```

```
F:189-219/Domain: MAP2/tau repeat homology <MT2>
F:220-251/Domain: MAP2/tau repeat homology <MT3>

Query Match 35.7%; Score 1415.5; DB 1; Length 316;
Best Local Similarity 74.5%; Pred. No. 5.1e-58;
Matches 286; Conservative 1; Mismatches 0; Indels 97; Gaps 2;

QY 375 KARMVSKSGDGTGDDKKAKTSTRSSAKTLKNRCLSPKLPFGSSDPLTPGSSPAVCPE 434
Db 30 QARVSKSGDGTGDDKKA----- 48
QY 435 PPSPKPVSSVTSRTSGGAKEMKLKGADGKTKIATPRGAAPPQCGQANATRIAPKTPP 494
Db 49 -----KGADGKTKIATPRGAAPPQCGQANATRIAPKTPP 83
QY 495 APKTPSSGPPKSGDRSGYSPGSPGTPGSRSTRPSLTPPTRPKKVAIVVTPPKSPS 554
Db 84 APKTPSSGPPKSGDRSGYSPGSPGTPGSRSTRPSLTPPTRPKKVAIVVTPPKSPS 143
QY 555 SAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKLQLSNVQSKGSKDNI 614
Db 144 SAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGK----- 180
QY 615 KHPVGGSGVQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVSEKLDKDRVQSKIGSL 674
Db 181 -----VQIVKPVLDLSKVTSCGSLGNIHHKPGGQVEVSEKLDKDRVQSKIGSL 232
QY 675 DNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVVS GDTSPRHLSNVSTGSI 734
Db 233 DNITHVPGGNGKKIETHKLTFRENAKAKTDHGAEIVYKSPVVS GDTSPRHLSNVSTGSI 292
QY 735 DMVDSPOLATLADEVSAKQGL 758
Db 293 DMVDSPOLATLADEVSAKQGL 316

RESULT 9
B28820
microtubule-associated protein tau type 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: B28820
R:Lee, G.; Cowan, N.; Kirschner, M.
Science 239, 285-288, 1988
A:Title: The primary structure and heterogeneity of tau protein from mouse brain.
A:Reference number: A94298; MUID:88099510; PMID:3122323
A:Accession: B28820
A:Molecule type: mRNA
A:Residues: 1-341 <LEE>
A:Cross-References: GB:M18775; NID:g201114; PIDN:AAA40165.1; PID:g201115
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; tandem repeat
F:183-213/Domain: MAP2/tau repeat homology <MT1>
F:214-244/Domain: MAP2/tau repeat homology <MT2>
F:245-276/Domain: MAP2/tau repeat homology <MT3>

Query Match 34.2%; Score 1354.5; DB 2; Length 341;
Best Local Similarity 41.2%; Pred. No. 3.3e-55;
Matches 313; Conservative 13; Mismatches 13; Indels 421; Gaps 7;

QY 1 MAEPQEFYEDHAGTYGIGDRKDGQGYTMHQDQEGDTDAGLKESPLQPTTBDGSEEPG 60
Db 1 MADPRQEFDTMEDHA-----GDYTLQDQEGMDHGLK----- 33
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAREAGIGDTPSLEDAAG 120
Db 34 -----AEEAGIGDTPNQEDQAAG 51
QY 121 HVTQEPESKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDLHOEGPLKAGGERPGSKKEVEDRDVDESSPQSPSKASPA 240
```



```
Db 56 ----- 55
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLPVDFLSKVSTETPASEPDGPSVGRAGQDAPLE 300
Db 56 ----- 59
QY 301 FTFHVEITPNVQKEAHSEHLGRAAFPAGPGEPEARGLSLEDTKADLPPEPSEKQPA 360
Db 60 ----- 59
QY 361 AAPRGKPVSRVQPKARVMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 60 ----- 71
QY 421 DPLIQSSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 72 -----
QY 479 QKQANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGTSGSRSTPSLTPPTPR 538
Db 93 QKTSNATRIPAKTPSPKTPPGSGEPKSGERSGYSSPGTSGSRSTPSLTPPTPR 152
QY 539 EPKKVAVVRTPPKSPSASRSRQTAPVPMPLKKNVSKIGSTENLKHQGGGKVQIINKK 598
Db 153 EPKKVAVVRTPPKSPSASRSRQTAPVPMPLKKNVSKIGSTENLKHQGGGK----- 205
QY 599 LDLSNVQSKGSKDNIKHVPGGGSGVQIYKPVDSLKVTSKCSLGNIIHHKPGGGOVEVKS 658
Db 206 -----
QY 659 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 718
Db 242 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 301

RESULT 10
A28820
microtubule-associated protein tau type 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: A28820
R:Lee, G.; Cowan, N.; Kirschner, M.
Science 239, 285-288, 1988
A:Title: The primary structure and heterogeneity of tau protein from mouse brain.
A:Reference number: A94298; MUID:88099510; PMID:312323
A:Accession: A28820
A:Molecule type: mRNA
A:Residues: 1-364 <LEE>
A:Cross-references: GB:M18776; NID:g201116; PIDN:AA40166.1; PID:g201117
C:Superfamily: microtubule-associated protein tau; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; tandem repeat
F:183-213/Domain: MAP2/tau repeat homology <MT1>
F:214-244/Domain: MAP2/tau repeat homology <MT2>
F:245-276/Domain: MAP2/tau repeat homology <MT3>

Query Watch 34.1%; Score 1350.5; DB 2; Length 364;
Best Local Similarity 41.1%; Pred. No. 5.4e-55;
Matches 312; Conservative 13; Mismatches 13; Indels 421; Gaps 7;

QY 1 MAPRPFEEFVMEHAGTYGLGDRKDGQGYTMHQDQEGDTHDGLKSPLOTPTEDGSEEPG 60
Db 1 MADPRQEFDTMEDHA-----GDYTLQLQDQEGDMDHGLK----- 33
QY 61 SETSDAKSTPTABDVTAPLVDEGAPCKQAQAQPHTEIPGTTAAEEAGIGDTPSLEDAAG 120
Db 34 -----AEEAGIGDTPNQDQAAG 51
QY 121 HVTQPESGKVQEGFLREPFPGLSHQLMSGNPGAPLLPEGPREATRQPSGTGPDTEG 180
```

```
Db 52 HVTQ----- 55
QY 181 GRHAPELLKHQLLDGLHQEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
Db 56 ----- 55
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLPVDFLSKVSTETPASEPDGPSVGRAGQDAPLE 300
Db 56 ----- 59
QY 301 FTFHVEITPNVQKEAHSEHLGRAAFPAGPGEPEARGLSLEDTKADLPPEPSEKQPA 360
Db 60 ----- 59
QY 361 AAPRGKPVSRVQPKARVMVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 60 ----- 71
QY 421 DPLIQSSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGADGKT--KIATPRGAAPP 478
Db 72 -----
QY 479 QKQANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGTSGSRSTPSLTPPTPR 538
Db 93 QKTSNATRIPAKTPSPKTPPGSGEPKSGERSGYSSPGTSGSRSTPSLTPPTPR 152
QY 539 EPKKVAVVRTPPKSPSASRSRQTAPVPMPLKKNVSKIGSTENLKHQGGGKVQIINKK 598
Db 153 EPKKVAVVRTPPKSPSASRSRQTAPVPMPLKKNVSKIGSTENLKHQGGGK----- 205
QY 599 LDLSNVQSKGSKDNIKHVPGGGSGVQIYKPVDSLKVTSKCSLGNIIHHKPGGGOVEVKS 658
Db 206 -----
QY 659 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 718
Db 242 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVS 301

RESULT 11
QRHUMT
microtubule-associated protein 2, splice form MAP-2b - human
N:Alternate names: MAP2
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000
C:Accession: I53693; A61085; PL0024; S34131
R:Albala, J.S.; Kalcheva, N.; Shafit-Zagardo, B.
Gene 136, 377-378, 1993
A:Title: Characterization of the transcripts encoding two isoforms of human microtubu
A:Reference number: I53693; MUID:94124038; PMID:8294038
A:Accession: I53693
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1824 <ALB>
A:Cross-references: EMBL:Z21958; GB:L12563; NID:g1850616; PIDN:AAB48098.1; PID:g18506
R:Dammerman, M.; Yen, S.H.; Shafit-Zagardo, B.
J. Neurosci. Res. 24, 487-495, 1989
A:Title: Sequence of a human MAP-2 region sharing epitopes with Alzheimer neurofibril
A:Reference number: A61085; MUID:90096190; PMID:2481044
A:Accession: A61085
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 77-645 <DAM>
R:Kosik, K.S.; Orecchio, L.D.; Bakalis, S.; Duffy, L.; Neve, R.L.
J. Neurochem. 51, 587-598, 1988
A:Title: Partial sequence of MAP2 in the region of a shared epitope with Alzheimer ne
A:Reference number: PL0024; MUID:88274407; PMID:2455776
A:Accession: PL0024
```


A:Molecule type: mRNA
A:Residues: 489-1538 <KOS>
A:Cross-references: GB:M25668; NID:g187380; PIDN:AAA59552.1; PID:g187381
C:Comment: Microtubule-associated proteins are a complex group consisting of the high molecular weight proteins that they may stabilize the microtubules against depolymerization.
C:Genetics:
A:Gene: GDB:MAP2
A:Cross-references: GDB:l18836; OMIM:157130
A:Map position: 2q34-2q35
C:Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
C:Keywords: alternative splicing; microtubule binding; phosphoprotein; tandem repeat
F:1455-1463/Region: microtubule binding #status predicted
F:1666-1696/Domain: MAP2/tau repeat homology <MT1>
F:1697-1727/Domain: MAP2/tau repeat homology <MT2>
F:1728-1759/Domain: MAP2/tau repeat homology <MT3>
F:657,958,1064,1250,1436,1503/Binding site: phosphate (Thr) (covalent) #status predicted
F:697,817,829,1320,1417,1542,1551/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 20.4%; Score 812.5; DB 1; Length 1824;
Best Local Similarity 29.2%; Pred. No. 1e-29;
Matches 284; Conservative 120; Mismatches 279; Indels 291; Gaps 39;

QY 5 ROFEVMEHAGTYGLGDRKDGKGYTMHODQEGDTDAGLKESPLQPTDGESEEPGSETS 64
Db 922 KDFSV-DKEASAHISGDK---SGUSKEFDQKKAN-----DRLDTVLEK-SEEHADSK 971
QY 65 DAKSTPTAEDVTAPLVDEGAPGKQAQAAPHTPEICTTAEEA--GIGDTPSLDEEAAGHV 122
Db 972 HAKKTEAGD-ELETFGLGVTEYEQALAK-DLSIPTDASSEKAEGKLSVPEI----- 1021
QY 123 TQPESEKVVQGF-----LREPGPGLSHQLMSGM-----PGAPLLPGPREA 166
Db 1022 -AEVPSKKVEQGLDFAVGQGLDVKISDFG-----QMASGLNIDRRATELKLKLEATQDM 1074
QY 167 TROPSTGTPEDT-----EGGRHAPELLKHOLLGD--LHQEGPPLK--GAGGK 209
Db 1075 T--PSSKAPQEDADAFMGVSGHMKETKVSETEVKQAKPDLVHQEAADVKEESYESSGE 1132
QY 210 ERPGSEEEVDEDRVDDESSPODS-----PP--SKASPAQDGRP----- 245
Db 1133 HESLTWESLKADGKKETSPESLIQDETAVKLSVEIPCPPAVSEADLATDERADVQMEF 1192
QY 246 ---PQTAAREATSIQGFPAEGAIPLPVDFLSKVSTEIPASEPDSVGRAGQDAPLFT 302
Db 1193 IOGPKESKETDISITPSDVAEPLHETIVSE-PAEIOSEEE---TEAAGEYDKLLFR 1247
QY 303 F-----HVEITPNVQK-----EQAHSEELHGRAAFPG 329
Db 1248 SDTLQITDLGVSGAREEFVETCPSEHKGVIESVVTIEDFTVVTQTTDDESGSHSVRF 1307
QY 330 APGEFPEARGPSIGEDTKREADLPPESEKO-----PAAAPRGKPVSRVPQLKARVMS 380
Db 1308 AALEQPEVERRRSPHDEEFEEVAAEAQAEPKDGSPAPASPEREEVA-LSEYKTEIYD 1366
QY 381 KSKDGTGSD-----KKATSTRSS-AKTLKNRP 408
Db 1367 DYKDETTIDDSIMDADSLWVDQDDRSIMTQLETIPKEEKAERARRSLEKHKRKEKP 1426
QY 409 C-----LSPKLPPTGSSDPL-----IOPSSPA--VCPPEP 436
Db 1427 FKTGRGRISTPERKVAKEPSTVSDEVRKKAIVYKKAELAKKTEVQAHSPSRKFILKPA 1486
QY 437 ---SSPKHVSYTSRTGSSGA-----KEMKLKADGKTKIATPRGAAP-----PG 478
Db 1487 IKYTRPThLSCVKRKTAAAGGESALAPSVEKQAKDKVSDGVTKSPKRSLSLPRSSILPP 1546
QY 479 QKG-----QANATRIKAPT-----PAKTPPPS-----GEPKSG 509
Db 1547 RRGVSGDRDENSFLNSISSARRTTRSEPIRRAGKSGTSTPTTPPGSTAITPGTPPPYS 1606
QY 510 DRSGVSGPGSGTSGRSRTPSLPTPP-----TRPKKAVAVVTRTPPKSPSSAKSLQATP 564
Db 1607 SR-----TPGTGTPP-SYPTPTPTGTPKSAILVPSEKKAIVIRTPPKSPGLTPKQLRLIN 1661

QY 565 VPMPLKNVSKSIGSTENLKHQPGGKGVQIINKLKLDSNVQSKGSKDNIKHVPGGGSVQ 624
Db 1662 QPLPDLKNVSKSIGSTDNIKYQPKGGQVIVTKKIDLSH----- 1700
QY 625 IVYVPVDSLKVTSKCGSLGNHHPKPGGQVEVSEKLDKDRVQSKIGSLDNTHVPPGGG 684
Db 1701 -----VTSKCGSLKNIRHRPGGGRVKIESVKLDFKEAKAQKVGSLDNAHHVPPGG 1750
QY 685 NKTIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDTPSRHLSNVSSVSTGSDMVDSPQAT 744
Db 1751 NVKIDSOKLAPREHAKARVDHGAIIITQSGRSSVASPRRLSNVSSSGSINLLESQLAT 1810
QY 745 LADEVASLAKQGL 758
Db 1811 LAEDVTAALAKQGL 1824

RESULT 12
A40115
microtubule-associated protein MAP2 - mouse
N:Alternate names: MAP2
C:Species: Mus musculus (house mouse)
C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 13-Aug-1999
C:Accession: A40115; S06467
R:Lewis, S.A.; Wang, D.; Cowan, N.J.
Science 242, 936-939, 1988
A:Title: Microtubule-associated protein MAP2 shares a microtubule binding motif with A:Reference number: A40115; MUID:89043973; PMID:3142041
A:Accession: A40115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1828 <LEW>
A:Cross-references: GB:M21041; NID:g199022; PIDN:AAA39490.1; PID:g199023
R:Wang, D.; Lewis, S.A.; Cowan, N.J.
Nucleic Acids Res. 16, 11369-11370, 1988
A:Title: Complete sequence of a cDNA encoding mouse MAP2.
A:Reference number: S06467; MUID:89083571; PMID:3205744
A:Accession: S06467
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-1828 <VAN>
A:Cross-references: GB:M21041; NID:g199022; PIDN:AAA39490.1; PID:g199023
C:Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
C:Keywords: microtubule binding; tandem repeat
F:1670-1700/Domain: MAP2/tau repeat homology <MT1>
F:1701-1731/Domain: MAP2/tau repeat homology <MT2>
F:1732-1763/Domain: MAP2/tau repeat homology <MT3>

Query Match 20.4%; Score 807; DB 2; Length 1828;
Best Local Similarity 28.3%; Pred. No. 1.8e-29;
Matches 266; Conservative 127; Mismatches 274; Indels 274; Gaps 33;

QY 3 EPROFEVMEHAGTYGLGDRKDGK-----YTMHQDQEGDTDAGLKESPLQPTDGESE- 57
Db 977 EHAKESSEMGKVELFGIGITYDQASTKEKLITTKDTSPKTEKGLSVPEVAEVEPTTKA 1036
QY 58 EPGETSDAKSTPTAEDVTAPLVDEGAPGKQA-----AAQPHTEIPECTTAAEAGIGDTPS 113
Db 1037 DQGLDPAATKAEPSSQDLIKVSDFGOMASGMNVNAGKAEIKFEVAQBELTLSSAPQADS 1096
QY 114 LEDBAAGHVTOEPESKVVQEGFLREPGPGLSHQLMSGMPGAPLLPEGPREATRPSGT 173
Db 1097 FMGVESGHI---KEGGKVNTEVKETKVPDLVHQ-----EAVDKESYESSG- 1141
QY 174 GPDETEGRHAPELLKHLQGLDLHQEGPPLKGAGGKRPCKSEKEVEDDRDVEDSSPODS- 232
Db 1142 -----EHEL-----TMESLKPDECKKETSPTSL 1166
QY 233 -----PPSKASPAQDGRP-----PQTAAREATSIQGFPAEGAIP 266
Db 1167 IQDEVALKLSVEIPCPPVSEADLSTDEKEVQMEFTQLPKREESTETPDIPAIPOVTOPT 1226

QY 267 LPVDFLSKVSTETIPASEPGSGVGRAGK-----QDAPLEFTF 303
Db 1227 QPEAIIVSE-PAEVPSEBEEIEAGGEYDKLLFRSDTLQISDLLVSEESREFEVETPCGELKG 1285
QY 304 HVEITPNVQ-----KEQAHSEHLGLRAAFPGAGGEPGPEARGPSLGEDTKADL-- 351
Db 1286 VVESVVTIEDFTTVQTTTDEGESGSHVRFA---APAQPEERPRPHDELETEMAA 1342
QY 352 ---PESEKOP---AAAPRGKPVSRVQLKARMVSKSGDGTGSD----- 390
Db 1343 EAQAEKPGSGDPAPATPEKEEVA-FSEYKTETDYDDYKDETTIDDSIMDADSLWDTQDDD 1401
QY 391 -----KRAKTSR---SSAKTLKNRPC-----LSPKLPTPGSSD 421
Db 1402 RSLTQLETIPKEERAEEKDARRPSLEKHKRKEPFKGTGRISTPERKVAKEPSTVSRD 1461
QY 422 PL-----IQSSPA---VCPEPP---SSPKHVSSVTSRTGSS----- 452
Db 1462 EVRRKKAVYKKAEALAKSEVQAHSPSRKLLKPAIKYTRTHLSVCVKRKTAAASGDLAQ 1521
QY 453 -GA-KEMKLAGDGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTPPAPKTP 499
Db 1522 PGAFKQAKQKVTGISKSPKRSLSRPPSSILPPRGVSGDRENSFSLNSSISSARRTT 1581
QY 500 PSSGPEPKSGDRSGYSSPGSGP---TPGS---RSRTPSLPTPPT---REP----- 540
Db 1582 RS--EPIRRAGKSGTSTPTTGGSTAITPGTPPSSYSSRTPGTCTGTPSYPRTPGTPKSGILV 1639
QY 541 ---KVAIVTTPKSPSSAKSRLOTPAPVMPDLKVNKSKIGSTENLKHQFGKQVLIIN 597
Db 1640 PSEKKVAIIRTPKSPATPK-QLRNLNQPLDLKVNKSKIGSTDNKYQPKGGQVQIVTK 1698
QY 598 KLDLSNVOSKCGSKDNKIKHPVGGSGVQIYKVPDLSKVTSCGSLGNIHKKPGGGOVEVK 657
Db 1699 KIDLSH-----VTSKCSLKNIRHPGGGRVKIE 1727
QY 658 SEKLDKDRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAIEVYKSPVVS 1717
Db 1728 SVKLDPEKAQKAVGSLDNAHHPVGGGNVYKIDSKLNFREHAKARVDHGAIEITQSPRS 1787
QY 718 GDTSPRHLNVSSGSDWVDSQPLATLADSVASLAKQGL 758
Db 1788 SVASPRLSNVSSGSLNLSQPLATLADSVTAALAKQGL 1828
RESULT 13
A37981
microtubule-associated protein 2b - rat
N;Alternate names: MAP2b
N;Contains: microtubule-associated protein 2c (MAP2c)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Jun-1991 #sequence_revision 20-Sep-1991 #text_change 13-Aug-1999
C;Accession: A37981; B37981; S10003; S07887; S14568
R;Kindler, S.; Schulz, B.; Goedert, M.; Garner, C.C.
J. Biol. Chem. 265, 19679-19684, 1990
A;Title: Molecular structure of microtubule-associated protein 2b and 2c from rat brain.
A;Reference number: A37981; MUID:91060576; PMID:2174050
A;Accession: A37981
A;Molecule type: mRNA
A;Residues: 1-1830 <KIN>
A;Cross-references: GB:X51842; NID:g56620; PIDN:CAA36135.1; PID:g56621
A;Accession: B37981
A;Molecule type: mRNA
A;Residues: 1-151,1515-1830 <K13>
R;Kindler, S.; Schwanke, B.; Garner, C.C.
Nucleic Acids Res. 18, 2822, 1990
A;Title: Complete cDNA sequence encoding rat high and low molecular weight MAP2.
A;Reference number: S10003; MUID:90251471; PMID:2339070
A;Accession: S10003
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-1830 <K12>
A;Cross-references: EMBL:X51842; NID:g56620; PIDN:CAA36135.1; PID:g56621

R;Doll, T.; Papandrikopoulou, A.; Matus, A.
Nucleic Acids Res. 18, 361, 1990
A;Title: Nucleotide and amino acid sequences of embryonic rat MAP2c.
A;Reference number: S07887; MUID:90221819; PMID:2326166
A;Accession: S07887
A;Molecule type: mRNA
A;Residues: 1-151,1515-1830 <DOI>
A;Cross-references: EMBL:X17682; NID:g56622; PIDN:CAA35667.1; PID:g56623
R;Matus, A.; Doll, T.
submitted to the EMBL Data Library, May 1990
A;Reference number: S14568
A;Accession: S14568
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-476,'H' 478-486,'E' 488-525,'R' 527-665,'V' 667-670,'K' 672-872,'R' 874
A;Cross-references: EMBL:X53455; NID:g57619; PIDN:CAA37535.1; PID:g57620
C;Genetics:
A;Gene: map2
C;Superfamily: microtubule-associated protein MAP2b; MAP2/tau repeat homology
C;Keywords: alternative splicing; microtubule binding; tandem repeat
F:1-1830/product: microtubule-associated protein 2b #status predicted <M2B>
F:1-151,1515-1830/product: microtubule-associated protein 2c #status predicted <M2C>
F:1672-1702/domain: MAP2/tau repeat homology <MT1>
F:1703-1733/domain: MAP2/tau repeat homology <MT2>
F:1734-1765/domain: MAP2/tau repeat homology <MT3>

Query Match 20.1%; Score 796.5; DB 2; Length 1830;

Best Local Similarity 29.6%; Pred. No. 5.5e-29;

Matches 276; Conservative 122; Mismatches 278; Indels 257; Gaps 35;

QY 3 EPRQFEVMEHAGTYGIG-----DRKQGGYTMHQEGGT---- 39
Db 978 EHAKESEVGDVLFGLGVYEQTSAKELITTKETAPERAEKGLSSVPEVAEVETTKA 1037
QY 40 DAGL-----KESPLQPTDEGSEEPGSETSDAKST-----PTAED- 74
Db 1038 DQGLDVAAKDDQSDLDIKVSDFGQWASGMSVDAGTKTELKPEVDQQLTSLSEAPQETDS 1097
QY 75 ---VTAPLVDECA-----PGKAAAQPHTEIPEGTTAEAA---GIGDTPSLEDEA---AG 120
Db 1098 FMGIESSHVKGDKGVSETEVKEKVPDLVHQEAVDKVEESYESSGEHESLTWESLKPDEG 1157
QY 121 HVTQPEPGKVVQEGFLR----EPGPPGLSHQLMS-----GMP-----GAPLL 159
Db 1158 KKETSPTSILQIDEVALKLSVEIPCPVPVSEADSSIDEKAEVQMEFIQLPREESTETPDI 1217
QY 160 PEGPREATR-QPSGTGPEDEGGHAPELLKHQLGLDHLQEGPPLKGAGGKER----- 211
Db 1218 PAIPSDVTPQPEAVVSEPAE-----VRGEETEAECEYDKLFRSDT 1261
QY 212 -----PGSKEEV-----DEDRVDDES---SPQDPPSKASPAQDGRPPOTAAREATS 255
Db 1262 LQITDLLVPGSREEFVETCPGHEKGVESVWTIEDFTTVQTTTDEGELGSHSVRFAA- 1320
QY 256 IPGFPAEGAIPVPVDFLSKV---STEIPASEPDGVSGRAGKGDADPLETTFHVEIIPNVQK 313
Db 1321 -PVQPEEERPPYDHELEVLMAEAQAEPKDGSPDAPATPEKEVPFSEYKTYETDYDK 1379
QY 314 EQAHSEHLGRA-----AFPGAPGEPGPEARGPSLGEDTKADLPE 353
Db 1380 DETTIDDSIMDADSLWDTQDDDRSILTQLETIPKEERAKEARRPSLEKHKR----- 1434
QY 354 PSEKQPAAPRGKPVSRVLPQLKARVSKSKDGTGS-DOKKAKTSTRSSAKTLKNRPCLS 412
Db 1435 ----KPFKTGRG-----RISTPERREVAKEPSTVSRDEVRKKKAVYKKAELAK----- 1479
QY 413 KLPTCGSSDPLIQSSPA--VCPEPP---SSPKHVSSVTSRT-----GSSGAKEMK 458
Db 1480 -----ESEVQAHSPSRKLLKPAIKYTRTHLSVCVKRKTATSGSAQAPSAFAKAK 1531
QY 459 LKGADGKTKIATPRGAAP-----PCQKG-----QANATRIPAKTPPAPKTPPSGEP 507
Db 1532 DKVTDGITKSPKRSLSLPRPSSILPPRGVSGDRENSFSLNSSISSARRTTRS--EPIR 1589


```
Db 75 INGEL--TSADRETAEEVSARIVQVVTAEAVAVLK-----GQEKEQAQHKDQTAAL----- 123
Qy 407 RCLSPKLTTPGSSDPLIQSPSPAVCPPEPPSPKHHVSSVTSTGSSGA-----KEMK 458
Db 124 -----PLAAETANLPPSPSPSPASEQVTVEEAAGGESALAPSVFKQAK 168
Qy 459 LKGADGKTKIATPRGAAP-----PCQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPEKRSLPRSSILPRRGVSGDRDENESFSLNSSISSSARRTTRSEPIRR 228
Qy 494 -----PAKTPPSS-----GEPKSGDRSGYSSPGSGTSGSRRTPSLTPPP-----TR 538
Db 229 ACKSGTSTPTPGSTAITPGTPSYSSR---TPGTPGTP-SYPRTPHTPGTPKSAIILVP 283
Qy 539 EPKKAIVVTPPKSPSSAKSRLQTAAPVMPDLKNVKSIGSTENLKHQPGGKVOIINKK 598
Db 284 SEKKAIIRTPPKSPGLTPKQLRLINQPLDKNVKSIGSTDNIKYQPGGQVQIVTKK 343
Qy 599 LDLSNVQSKCGSKDNIKHVPGGGVQIVYKPVDSLKVTSKGSLGNIHHKPGGGGOVEVKS 658
Db 344 IDLSH-----VTSKCSLKNIRHPPGGGRVKIES 372
Qy 659 EKLDFKDRVQSKIGSLDNITHVPGGKNKIEHTKLITFRENAKAKTDHGAEIYKSPVVYSG 718
Db 373 VKLDFEKAQAQKVGSLDNAHHVPGGKNKIDSKLNFREHAKARVDHGAEIITQSPGRSS 432
Qy 719 DTSPRHLSNVSTGSDWVDSPLATLADEVSAKQGL 758
Db 433 VASPRRLSNVSSSGSINLLESPLATLAEDVTAALAKQGL 472
```

Search completed: January 28, 2003, 14:20:12
Job time : 29.8765 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 11.5946 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-5
Perfect score: 3965
Sequence: 1 MAEPQFEFVEMDHAGTYGL.....SPQLATLADEVSAKLQGL 758

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3960	99.9	757	1 TAU_HUMAN	P10636 homo sapien
2	2852	71.9	751	1 TAU_RAT	P19332 rattus norv
3	2782.5	70.2	732	1 TAU_MOUSE	P10637 mus musculu
4	2032.5	51.3	458	1 TAU_MACMU	P57786 macaca mula
5	1840	46.4	447	1 TAU_BOVIN	P29172 bos taurus
6	1744.5	44.0	382	1 TAU_PAPHA	Q96VX8 papio hamad
7	1711.5	43.2	402	1 TAU_CAPIH	O02828 capra hircu
8	937	23.6	1861	1 MAP2_RAT	P15146 rattus norv
9	815	20.6	1827	1 MAP2_HUMAN	P11137 homo sapien
10	807	20.4	1828	1 MAP2_MOUSE	P20357 mus musculu
11	624	15.7	1152	1 MAP4_HUMAN	P27816 homo sapien
12	621.5	15.7	1125	1 MAP4_MOUSE	P27546 mus musculu
13	545	13.7	1072	1 MAP4_BOVIN	P36225 bos taurus
14	297	7.5	1411	1 TCOF_HUMAN	Q13428 homo sapien
15	294	7.4	1464	1 CAL3_MOUSE	P08121 mus musculu
16	290	7.3	1049	1 CAL3_BOVIN	P04258 bos taurus
17	287.5	7.3	1806	1 CALB_HUMAN	P12107 homo sapien
18	283.5	7.2	779	1 CALB_BOVIN	P02453 bos taurus
19	283.5	7.2	1466	1 CAL3_HUMAN	P02461 homo sapien
20	283	7.1	1185	1 DRPL_HUMAN	P54259 homo sapien
21	282	7.1	1262	1 CAL3_CHICK	P12105 gallus gall
22	279.5	7.0	1685	1 CA54_HUMAN	P29400 homo sapien
23	278.5	7.0	1453	1 CAL1_CHICK	P20457 gallus gall
24	278.5	7.0	1804	1 CALB_MOUSE	O61245 mus musculu
25	277.5	7.0	1763	1 CA24_ASCSU	P27393 ascaris suu
26	275	6.9	911	1 CALB_BOVIN	Q28083 bos taurus
27	273.5	6.9	1210	1 AF4_HUMAN	P51825 homo sapien
28	273	6.9	1838	1 CAL5_HUMAN	P20908 homo sapien
29	271	6.8	1459	1 CAL2_MOUSE	P28481 mus musculu
30	268.5	6.8	1736	1 CA2B_HUMAN	P13942 homo sapien
31	268	6.8	1183	1 DRPL_RAT	P54258 rattus norv
32	268	6.8	2944	1 CAL7_HUMAN	Q02388 homo sapien
33	267.5	6.7	660	1 YHLL_EBV	P03181 epstein-bar

ALIGNMENTS

RESULT 1

ID	TAU_HUMAN	STANDARD;	PRT;	757 AA.
AC	P10636; P18518; Q14799; Q15551; Q9UQ96; Q15549; Q15550; Q9UDJ3;			
AC	Q9UMH0;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Microtubule-associated protein tau (Neurofibrillary tangle protein)			
DE	(Paired helical filament-tau) (PHF-tau).			
GN	MAPT OR MTBT1 OR TAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]				
RP	SEQUENCE FROM N.A. (ISOFORMS PNS-TAU; TAU-A AND TAU-F).			
RA	Andreadis A.;			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
[2]				
RP	SEQUENCE FROM N.A. (ISOFORM TAU-A).			
RC	TISSUE=Brain;			
RX	MEDLINE=88234557; PubMed=3131773;			
RA	Goedert M., Wischik C., Crowther R., Walker J., Klug A.;			
RT	"Cloning and sequencing of the cDNA encoding a core protein of the			
RT	paired helical filament of Alzheimer disease: identification as the			
RT	microtubule-associated protein tau.;			
RL	proc. Natl. Acad. Sci. U.S.A. 85:4051-4055(1988).			
[3]				
RP	SEQUENCE FROM N.A. (ISOFORMS TAU-B; TAU-C; TAU-E AND TAU-F).			
RC	TISSUE=Brain;			
RX	MEDLINE=90380393; PubMed=2484340;			
RA	Goedert M., Spillantini M.G., Jakes R., Rutherford D., Crowther R.A.;			
RT	"Multiple isoforms of human microtubule-associated protein tau:			
RT	sequences and localization in neurofibrillary tangles of Alzheimer's			
RT	disease.;			
RL	Neuron 3:519-526(1989).			
[4]				
RP	SEQUENCE FROM N.A. (ISOFORM TAU-D).			
RC	TISSUE=Brain;			
RX	MEDLINE=89251564; PubMed=2498079;			
RA	Goedert M., Spillantini M.G., Potier M.C., Ulrich J., Crowther R.A.;			
RT	"Cloning and sequencing of the cDNA encoding an isoform of			
RT	microtubule-associated protein tau containing four tandem repeats:			
RT	differential expression of tau protein mRNAs in human brain.;			
RL	EMBO J. 8:393-399(1989).			
[5]				
RP	SEQUENCE FROM N.A. (ISOFORMS TAU-A AND FETAL-TAU).			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=90180482; PubMed=2516729;			
RA	Lee G., Neve R.L., Kosik K.S.;			
RT	"The microtubule binding domain of tau protein.;"			
RL	Neuron 2:1615-1624(1989).			
[6]				
RP	SEQUENCE FROM N.A. (ISOFORM TAU-F), AND ALTERNATIVE SPLICING.			
RX	MEDLINE=93041757; PubMed=1420178;			

34	267.5	6.7	1650	1 CA2B_MOUSE	Q64739 mus musculu
35	265	6.7	1464	1 CAL1_HUMAN	P02452 homo sapien
36	265	6.7	1669	1 CAL4_MOUSE	P02463 mus musculu
37	264	6.7	1453	1 CAL1_MOUSE	P11087 mus musculu
38	264	6.7	1460	1 CAL1_CANFA	Q9XJ17 canis famil
39	263.5	6.6	1603	1 CALF_HUMAN	Q07092 homo sapien
40	263	6.6	1581	1 PPRB_HUMAN	Q15648 h peroxisom
41	262.5	6.6	674	1 CALA_BOVIN	P32006 bos taurus
42	261.5	6.6	671	1 CALA_RAT	P02454 rattus norv
43	261.5	6.6	747	1 CAL2_BOVIN	P02459 bos taurus
44	261	6.6	680	1 CALA_HUMAN	Q03692 homo sapien
45	261	6.6	1418	1 CAL2_HUMAN	P02458 homo sapien

- RA Andreadis A., Brown W.M., Kosik K.S.;
 RT "Structure and novel exons of the human tau gene.";
 RL Biochemistry 31:10626-10633(1992).
 RN [7]
 RN SEQUENCE OF 591-621 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=89193714; PubMed=2495000;
 RA Mori H., Hamada Y., Kawaguchi M., Honda T., Kondo J., Ihara Y.;
 RT "A distinct form of tau is selectively incorporated into Alzheimer's
 RT paired helical filaments.";
 RL Biochem. Biophys. Res. Commun. 159:1221-1226(1989).
 RN [8]
 RN SEQUENCE OF 1-72; 102-380; 467-496; 507-570; 576-582; 591-606;
 RP 615-633; 638-656; 660-663; 670-699 AND 702-757.
 RC TISSUE=Brain;
 RX MEDLINE=92381012; PubMed=1512244;
 RA Hasegawa M., Morishima-Kawashima M., Takio K., Suzuki M., Titani K.,
 RA Ihara Y.;
 RT "Protein sequence and mass spectrometric analyses of tau in the
 RT Alzheimer's disease brain.";
 RL J. Biol. Chem. 267:17047-17054(1992).
 RN [9]
 RN SEQUENCE OF 576-583; 607-610; 615-627; 638-647 AND 670-685,
 RP PHOSPHORYLATION, AND MUTAGENESIS.
 RX MEDLINE=95221434; PubMed=7706316;
 RA Drewes G., Trinczek B., Illenberger S., Biernat J., Schmitt-Ulms G.,
 RA Meyer H.E., Mandelkow E.-M., Mandelkow E.;
 RT "Microtubule-associated protein/microtubule affinity-regulating kinase
 RT interactions and dynamic instability that regulates tau-microtubule
 RT Alzheimer-specific site serine 262.";
 RL J. Biol. Chem. 270:7679-7688(1995).
 RN [10]
 RN REVIEW.
 RX MEDLINE=91320377; PubMed=1713721;
 RA Goedert M., Crowther R.A., Garner C.C.;
 RT "Molecular characterization of microtubule-associated proteins tau and
 RT MAP2.";
 RL Trends Neurosci. 14:193-199(1991).
 RN [11]
 RN SUBCELLULAR LOCATION, AND PHOSPHORYLATION.
 RX MEDLINE=20283597; PubMed=10747907;
 RA Maas T., Eidenmueller J., Brandt R.;
 RT "Interaction of tau with the neural membrane cortex is regulated by
 RT phosphorylation at sites that are modified in paired helical
 RT filaments.";
 RL J. Biol. Chem. 275:15733-15740(2000).
 RN [12]
 RN PHOSPHORYLATION, AND MUTAGENESIS.
 RX MEDLINE=98413833; PubMed=9735171;
 RA Sengupta A., Kabat J., Novak M., Wu Q., Grundke-Iqbal I., Iqbal K.;
 RT "Phosphorylation of tau at both Thr 231 and Ser 262 is required for
 RT maximal inhibition of its binding to microtubules.";
 RL Arch. Biochem. Biophys. 357:299-309(1998).
 RN [13]
 RN PHOSPHORYLATION, AND MUTAGENESIS.
 RX MEDLINE=98278959; PubMed=9614189;
 RA Illenberger S., Zheng-Fischer O., Preuss U., Stamer K., Baumann K.,
 RA Trinczek B., Biernat J., Godemann R., Mandelkow E.-M., Mandelkow E.;
 RT "The endogenous and cell cycle-dependent phosphorylation of tau
 RT protein in living cells: implications for Alzheimer's disease.";
 RL Mol. Biol. Cell 9:1495-1512(1998).
 RN [14]
 RN GLYCATION.
 RX MEDLINE=97465580; PubMed=9326300;
 RA Nacharaju P., Ko L., Yen S.H.;
 RT "Characterization of in vitro glycation sites of tau.";
 RL J. Neurochem. 69:1709-1719(1997).
 RN [15]
 RN REVIEW ON VARIANTS.
 RX MEDLINE=20437008; PubMed=10899436;
 RA Goedert M., Spillantini M.G.;
 RT "Tau mutations in frontotemporal dementia FTDP-17 and their relevance
 for Alzheimer's disease.";
 RL Biochim. Biophys. Acta 1502:110-121(2000).
 RN [16]
 RN VARIANT FTDP17 M-653, AND VARIANTS N-284; A-288; Y-440 AND P-446.
 RX MEDLINE=98291804; PubMed=9629852;
 RA Poorkaj P., Bird T.D., Wijsman E., Nemens E., Garruto R.M.,
 RA Anderson L., Andreadis A., Wiederholt W.C., Raskind M.,
 RA Schellenberg G.D.;
 RT "Tau is a candidate gene for chromosome 17 frontotemporal dementia.";
 RL Ann. Neurol. 43:815-825(1998).
 RN [17]
 RN ERRATUM.
 RP Poorkaj P., Bird T.D., Wijsman E., Nemens E., Garruto R.M.,
 RA Anderson L., Andreadis A., Wiederholt W.C., Raskind M.,
 RA Schellenberg G.D.;
 RL Ann. Neurol. 44:428-428(1998).
 RN [18]
 RN VARIANT FTDP17 LEU-617.
 RX MEDLINE=98409513; PubMed=9736786;
 RA Dumanchin C., Camuzat A., Campion D., Verpillat P., Hannequin D.,
 RA Dubois B., Saugier-Verber P., Martin C., Penet C., Charbonnier F.,
 RA Agid Y., Frebourg T., Brice A.;
 RT "Segregation of a missense mutation in the microtubule-associated
 RT protein tau gene with familial frontotemporal dementia and
 RT parkinsonism.";
 RL Hum. Mol. Genet. 7:1825-1829(1998).
 RN [19]
 RN VARIANTS FTDP17 VAL-588; LEU-617 AND TRP-722.
 RX MEDLINE=98303385; PubMed=9641683;
 RA Hutton M., Lendon C.L., Rizzu P., Baker M., Froelich S., Houlden H.,
 RA Pickering-Brown S., Chakraverty S., Isaacs A., Grover A., Hackett J.,
 RA Adamson J., Lincoln S., Dickson D., Davies P., Petersen R.C.,
 RA Stevens M., de Graaff E., Wauters E., van Baren J., Hillebrand M.,
 RA Joosse M., Kwon J.M., Nowotny P., Che L.K., Norton J., Morris J.C.,
 RA Reed L.A., Trojanowski J., Basun H., Lannfelt L., Neystat M., Fahn S.,
 RA Dark F., Tannenberg T., Dodd P.R., Hayward N., Kwok J.B.J.,
 RA Schofield P.R., Andreadis A., Snowden J., Craufurd D., Neary D.,
 RA Owen F., Oostra B.A., Hardy J., Goate A., van Swieten J., Mann D.,
 RA Lynch T., Heutink P.;
 RT "Association of missense and 5'-splice-site mutations in tau with the
 RT inherited dementia FTDP-17.";
 RL Nature 393:702-705(1998).
 RN [20]
 RN VARIANT PPND LYS-595, AND VARIANT FTDP17 LEU-617.
 RX MEDLINE=99007274; PubMed=9789048;
 RA Clark L.N., Poorkaj P., Wszolek Z., Geschwind D.H., Nasreddine Z.S.,
 RA Miller B., Li D., Payami H., Awert F., Markopoulou K., Andreadis A.,
 RA D'Souza I., Lee V.M.-Y., Reed L., Trojanowski J.Q., Zhukareva V.,
 RA Bird T., Schellenberg G., Wilhelmsen K.C.;
 RT "Pathogenic implications of mutations in the tau gene in
 RT pallido-ponto-nigral degeneration and related neurodegenerative
 RT disorders linked to chromosome 17.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13103-13107(1998).
 RN [21]
 RN VARIANTS FTDP17 VAL-588; LYS-596 DEL; LEU-617 AND TRP-722.
 RX MEDLINE=99138654; PubMed=9973279;
 RA Rizzu P., Van Swieten J.C., Joosse M., Hasegawa M., Stevens M.,
 RA Tibben A., Niermeijer M.F., Hillebrand M., Ravid R., Oostra B.A.,
 RA Goedert M., van Dulijn C.M., Heutink P.;
 RT "High prevalence of mutations in the microtubule-associated protein
 RT tau in a population study of frontotemporal dementia in the
 RT Netherlands.";
 RL Am. J. Hum. Genet. 64:414-421(1999).
 RN [22]
 RN VARIANTS FTDP17 LEU-617; MET-653 AND TRP-722.
 RX MEDLINE=99229757; PubMed=10214944;
 RA Nacharaju P., Lewis J., Easson C., Yen S., Hackett J., Hutton M.,
 RA Yen S.H.;
 RT "Accelerated filament formation from tau protein with specific FTDP-17
 RT missense mutations.";
 RL FEBS Lett. 447:195-199(1999).
 RN [23]
 RN VARIANT FTDP17/CBD SER-617.

Query Match 99.9%; Score 3960; DB 1; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.2e-144;
Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AEPQREVEFVMDHAGTYGLGDRKDDGGYTMHQDEGDTAGLKESPLQTPTDSEEPGS 61
DB 1 AEPQREVEFVMDHAGTYGLGDRKDDGGYTMHQDEGDTAGLKESPLQTPTDSEEPGS 60

QY 62 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEGTABEAGIGDTPTSEDAAGH 121
DB 61 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEGTABEAGIGDTPTSEDAAGH 120

QY 122 VTQPESSKVVQEGFLREPGLSHQMSGMPGAPLLPGPREATQPSGTGPEDEGG 181
DB 121 VTQPESSKVVQEGFLREPGLSHQMSGMPGAPLLPGPREATQPSGTGPEDEGG 180

QY 182 RHAPELLKHQLGLDHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQ 241
DB 181 RHAPELLKHQLGLDHOEGPPLKAGGKERPGSKKEEVEDRDVDESSPQSPSKASPAQ 240

QY 242 DGRPQTAAREATSIQCPFAEGAIPLVDFLSKVSTEIPASEPDGVSVGRAKQDAPLEF 301
DB 241 DGRPQTAAREATSIQCPFAEGAIPLVDFLSKVSTEIPASEPDGVSVGRAKQDAPLEF 300

QY 302 TFHVEITPNVQKEAHSEEHGRLAAPPAGPEGRPSLGEDTKKADLPPESEKQPA 361
DB 301 TFHVEITPNVQKEAHSEEHGRLAAPPAGPEGRPSLGEDTKKADLPPESEKQPA 360

QY 362 APRGKPVSRVPLKARVSKSDGTGDDKAKTSTRSSAKTLKRPCLSPKLPPTPGSSD 421
DB 361 APRGKPVSRVPLKARVSKSDGTGDDKAKTSTRSSAKTLKRPCLSPKLPPTPGSSD 420

QY 422 PLIQSSPAPCPPEPPSPKHSVTSRTGSSGAKEMKLGADCKTIATPRGAAPCGQK 481
DB 421 PLIQSSPAPCPPEPPSPKHSVTSRTGSSGAKEMKLGADCKTIATPRGAAPCGQK 480

QY 482 QANATRIAPKTPAPKTPPPSSGEGPPKSGDRSGYSSPGSGTSGRSRTSLPTPTPREK 541
DB 481 QANATRIAPKTPAPKTPPPSSGEGPPKSGDRSGYSSPGSGTSGRSRTSLPTPTPREK 540

QY 542 KVAVRTPPKSPSASRLQATVPMPDLKNVSKIGSTENLKHQGGGVQVINKKLDL 601
DB 541 KVAVRTPPKSPSASRLQATVPMPDLKNVSKIGSTENLKHQGGGVQVINKKLDL 600

QY 602 SNVQKCGSKDNKHVPGGSGVQIVKPYDLSKVTSCGSLGNHKKPGGGQVEVSEKL 661
DB 601 SNVQKCGSKDNKHVPGGSGVQIVKPYDLSKVTSCGSLGNHKKPGGGQVEVSEKL 660

QY 662 DFDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDTS 721
DB 661 DFDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDTS 720

QY 722 PRHLSNVSTGSDIMVDSPOLATLADVSASLAKQGL 758
DB 721 PRHLSNVSTGSDIMVDSPOLATLADVSASLAKQGL 757

RESULT 2

TAU_RAT
ID TAU_RAT STANDARD; PRT; 751 AA.
AC P19332; Q63567; Q9QW06; Q63577;
DT 01-NOV-1990 (Rel. 16, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
DE (paired helical filament-tau) (PHF-tau).
GN MAPT OR MTAPO OR TAU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]

SEQUENCE FROM N.A. (ISOFORM TAU-B).
TISSUE=Pheochromocytoma;
MEDLINE=92179305; PubMed=1542696;
RA Goedert M., Spillantini M.G., Crowther R.A.;
RT "Cloning of a big tau microtubule-associated protein characteristic of
RL the peripheral nervous system.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:1983-1987(1992).

SEQUENCE FROM N.A. (ISOFORM TAU-B).
TISSUE=Dorsal root ganglion;
MEDLINE=94013081; PubMed=8408300;
RA Georgieff I.S., Liem R.K.H., Couchie D., Mavillia C., Nunez J.,
Shelanski M.L.;
RT "Expression of high molecular weight tau in the central and peripheral
RL nervous systems.";
RN J. Cell Sci. 105:729-737(1993).

SEQUENCE FROM N.A. (ISOFORM TAU-F).
STRAIN=Wistar; TISSUE=Brain;
MEDLINE=94334997; PubMed=8057376;
RA Sadot E., Marx R., Barg J., Behar L., Ginzburg I.;
RT "Complete sequence of 3'-untranslated region of tau from rat central
RL nervous system. Implications for mRNA heterogeneity.";
RN J. Mol. Biol. 241:325-331(1994).

SEQUENCE FROM N.A. (ISOFORMS TAU-E AND TAU-G).
TISSUE=Brain;
MEDLINE=90180457; PubMed=2560640;
RA Kosik K.S., Orecchio L.D., Bakalis S., Neve R.L.;
RT "Developmentally regulated expression of specific tau sequences.";
RN Neuron 2:1389-1397(1989).

SEQUENCE FROM N.A. (ISOFORMS TAU-E AND TAU-C).
MEDLINE=89359509; PubMed=2504728;
RA Kanai Y., Takemura R., Oshima T., Mori H., Ihara Y., Yanagisawa M.,
Masaki T., Hirokawa N.;
RT "Expression of multiple tau isoforms and microtubule bundle formation
RL in fibroblasts transfected with a single tau cDNA.";
RN J. Cell Biol. 109:1173-1184(1989).

SEQUENCE OF 359-460 FROM N.A. (ISOFORM TAU-A), AND SEQUENCE OF 105-112
AND 367-460 FROM N.A. (ISOFORM TAU-D).
TISSUE=Spinal cord;
MEDLINE=95054048; PubMed=7964751;
RA Mavillia C., Couchie D., Nunez J.;
RT "Diversity of high-molecular-weight tau proteins in different regions
RL of the nervous system.";
RN J. Neurochem. 63:2300-2306(1994).

SEQUENCE OF 696-751 FROM N.A. (ISOFORMS TAU-A TO TAU-G), AND SEQUENCE
OF 751-774 FROM N.A. (ISOFORM TAU-H).
STRAIN=Sprague-Dawley; TISSUE=Brain;
MEDLINE=95182802; PubMed=7877441;
RA Sawa A., Oyama F., Matsushita M., Ihara Y.;
RT "Molecular diversity at the carboxyl terminus of human and rat tau.";
RN Brain Res. Mol. Brain Res. 27:111-117(1994).

-1- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
-1- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS. IN THE
CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
-1- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; TAU-A/SC1 (SHOWN HERE),
TAU-B/BIG-TAU/HMW-TAU, TAU-C, TAU-D/SC2, TAU-E, TAU-F, TAU-
G/FETAL-TAU AND TAU-H; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY
DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE OF UP TO 4 OF
THE 14 EXONS. TWO DIFFERENT C-TERMINI ARE OBTAINED EITHER BY THE
RETENTION OR THE SPLICING OF INTRON 13/14.

RP SEQUENCE FROM N.A. (ISOFORMS TAU-A; TAU-D AND TAU-E).
 RC STRAIN-Him OF1; TISSUE=Liver, Kidney, and Brain;
 RX MEDLINE=95012085; PubMed=797211;
 RA Kenner L., el-Shabravi Y., Hutter H., Forstner M., Zatloukal K.,
 RA Hoefler G., Preisegger K.-H., Kurbauer R., Denk H.;
 RT "Expression of three- and four-repeat tau isoforms in mouse liver.";
 RL Hepatology 20:1086-1089(1994).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORMS TAU-B AND TAU-C).
 RP TISSUE=Brain;
 RC MEDLINE=88099510; PubMed=3122323;
 RX Lee G., Cowan N.J., Kirschner M.;
 RA "The primary structure and heterogeneity of tau protein from mouse
 RT brain.";
 RL Science 239:285-288(1988).
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM B).
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=95182802; PubMed=7877441;
 RA Sawa A., Oyama F., Matsushita M., Ihara Y.;
 RT "Molecular diversity at the carboxyl terminus of human and rat tau.";
 RL Brain Res. Mol. Brain Res. 27:111-117(1994).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=94005827; PubMed=8402267;
 RA Couchie D., Gache Y., Mavillia C., Guilleminot J., Bridoux A.-M.,
 RA Nivez M.-P., Nunez J.;
 RT "High molecular weight tau proteins and acquisition of neuronal
 RT polarity in peripheral nervous system.";
 RL C. R. Acad. Sci., III, Sci. Vie 316:404-409(1993).
 CC -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
 CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
 CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
 CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
 CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
 CC -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
 CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS; PNS-TAU (SHOWN HERE),
 CC TAU-A, TAU-B, TAU-C, TAU-D AND TAU-E; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THEY DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE
 CC OF UP TO 5 OF THE 14 EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS
 CC THE ADDITIONAL TAU/MAP REPEAT. TWO DIFFERENT C-TERMINI ARE
 CC OBTAINED EITHER BY THE RETENTION OR THE SPLICING OF INTRON 13/14.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS AND AT A LOWER LEVEL IN
 CC THE LIVER AND KIDNEY. PNS-TAU IS EXPRESSED IN THE PERIPHERAL
 CC NERVOUS SYSTEM WHILE THE OTHERS ARE EXPRESSED IN THE CENTRAL
 CC NERVOUS SYSTEM.
 CC -!- DEVELOPMENTAL STAGE: SHORTER FORMS OR LOW MOLECULAR WEIGHT TAU
 CC (LMW-TAU) ARE GENERALLY EXPRESSED AT EARLY DEVELOPMENT STAGES AND
 CC LONGER FORMS OR HIGH MOLECULAR WEIGHT TAU (HMW-TAU) IN THE ADULT
 CC BRAIN.
 CC -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
 CC CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
 CC -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
 CC CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
 CC MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
 CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
 CC -!- DISEASE: MAY BE INVOLVED IN THE PATHOGENESIS OF CYTOSPLASMIC
 CC INCLUSIONS (AS MALLORY BODIES) IN LIVERS OF MICE CHRONICALLY
 CC INTOXICATED WITH GRISOFLUVIN OR DDC (3,5-DIETHOXYCARBONYL-2,4-
 CC DIHYDROCOLIDINE), A MODEL FOR HUMAN ALCOHOLIC HEPATITIS.
 CC ALTERATION OF TAU (ABNORMAL PHOSPHORYLATION AND CROSSLINKING)
 CC COULD CONTRIBUTE TO MALLORY BODIES FORMATION AND CROSSLINKING)
 CC MICROTUBULE FUNCTION IN ALCOHOLIC LIVER DISEASE.
 CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U12914; AAA58343.1; -;
 DR EMBL; U12915; AAA58344.1; -;
 DR EMBL; U12916; AAA58345.1; -;
 DR EMBL; Z12133; CAA78121.1; -;
 DR EMBL; M93266; -; NOT_ANNOTATED_CDS.
 DR EMBL; M18775; AAA40165.1; -;
 DR EMBL; M18776; AAA40166.1; -;
 DR EMBL; D30627; BAA18878.1; -;
 DR PIR; A28820; A28820.
 DR PIR; A28820; B28820.
 DR MGI; 97180; Mapt.
 DR InterPro; IPR002955; Tau_protein.
 DR InterPro; IPR01084; Tubulin_Tau.
 Pfam; PF00418; tubulin-binding; 4.
 PRINTS; PR01261; TAUPROTEIN.
 DR PROSITE; PS00229; TAU_MAP; 4.
 KW Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
 KW Phosphorylation.
 FT INIT_MET 0
 FT REPEAT 535 565
 FT REPEAT 566 596
 FT REPEAT 597 627
 FT REPEAT 628 659
 FT MOD_RES 1 1
 FT DISULFID 582 613
 FT VARSPLIC 33 90
 FT VARSPLIC 91 112
 FT VARSPLIC 113 349
 FT VARSPLIC 367 432
 FT VARSPLIC 566 596
 FT VARSPLIC 732 732
 FT CONFLICT 2 2
 FT CONFLICT 8 8
 FT CONFLICT 527 527
 FT CONFLICT 671 671
 SQ SEQUENCE 732 AA; 76112 MW; BFD0767E41C7A3A CRC64;
 BY SIMILARITY.
 TAU/MAP MOTIF 1.
 TAU/MAP MOTIF 2.
 TAU/MAP MOTIF 3.
 TAU/MAP MOTIF 4.
 ACETYLATION (BY SIMILARITY).
 BY SIMILARITY.
 MISSING (IN ISOFORM TAU-B, ISOFORM TAU-C,
 ISOFORM TAU-D AND ISOFORM TAU-E).
 MISSING (IN ISOFORM TAU-E).
 MISSING (IN ISOFORM TAU-A, ISOFORM TAU-B,
 ISOFORM TAU-C, ISOFORM TAU-D AND ISOFORM
 TAU-E).
 MISSING (IN ISOFORM TAU-A, ISOFORM TAU-B,
 ISOFORM TAU-C, ISOFORM TAU-D AND ISOFORM
 TAU-E).
 MISSING (IN ISOFORM TAU-B AND ISOFORM
 TAU-C).
 L -> KAALLSSOVNYSHDLATITDGL (IN
 ISOFORM TAU-B).
 D -> N (IN REF. 1).
 D -> N (IN REF. 1).
 P -> T (IN REF. 2; CAA78121).
 E -> Q (IN REF. 1).
 Query Match 70.28; Score 2782.5; DB 1; Length 732;
 Best Local Similarity 74.38; Pred. No. 1.2e-99;
 Matches 569; Conservative 45; Mismatches 109; Indels 43; Gaps 11;

Qy 2 AEPQREFVMDHAGTGLGDRKDGQGYTHQDGGTDAGLKESPLQTPTDGESEPGS 61
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 1 ADPQREFDTHDHA-----GDTLLQDQEGDMHGLKESPPQPPADGGAEEPGS 49
 Qy 62 ETSDAKSTPTAEDVTAPLVDEGAPCKAAQAQPHTEPGTTAEAGIDGTPSLEDEAAGH 121
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
 Db 50 ETSDAKSTPTAEDVTAPLVDERAPDKAAQAQPHTEPGTTAEAGIDGTPNQEDQAAGH 109
 Qy 122 VTQPESEKVVQEGFLREPFGPLSHQLMSGMPAPLLPEGPRAATQPSGTGPDTEGG 181
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 110 VTQ-----GR-----REGQAPDLGTSMDWTQQVSSMSGAPLLPQGLREATCQPSGTREDIEKS 163
 Qy 182 RHAPPELLKHQLGLDLHQBGLKAGKERKPSKEEVEDRDVDSESSPPSKASPAQ 241
 :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 164 HPASELLR-----RGPPQKEGWDRLGSEEEVEDEDTVDSESS-QDPPPSQSLTP 213
 Qy 242 DGRPPQ-----TAAEATSIPIGFAEAGIPLPVDFLSKVSFTEIPASEPDGPSVG-RAKGQD 296
 :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
 Db 214 GRAAPQAGSGSGVCATSVPLPTEGVSVPPLPADDFSKVSATQASQPEGCTGPMESGHE 273


```

QY 297 APLEFTFHFVEITPNVQEAHSEHLGRAAFPGAPGGEPA--RGPISLGEDTKADLP 354
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 274 AAEFTFHFVEIKASTPEQ-----DEGATVVGVPGEQKAQTQGPSVGKGTKEASLQEP 328
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 355 SEKQAAAAPRGKPVSRVPLKARMSKSDGTGDDKAKTSTRSAKTLKNRCLSPKL 414
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 329 PGKQPAAGLPGRPVSRVPLKARVA--SKDRGTGNDKKAKTSTPSCAKAPSHRPLSPTR 386
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 415 PTGSSDPLIOPSSPAVCPEPPSPKPHSVTSRTGSSCAKEMKLGADGKT--KIATPR 472
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 387 PTLGSSDPLIKPSSPAVSPATSPKPHSVTSVTPRNGSPCTKMKLGADGKCAKIATPR 446
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 473 GAAPPGKGQANATRIAPKTPPAPKTPPGSGEPKSGDGRSGYSSGSPGTPGSRSTRPSL 532
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 447 GAASPAQKGTSNATRIAPKTPPSPKTPPGSGEPKSGERSGYSSGSPGTPGSRSTRPSL 506
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 533 PTPPTREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGGKV 592
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 507 PTPPTREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGGKV 566
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 593 QIINKKLDLSNVQSKGSDKNIKHVPGGGQVQIVYKPVDSLKVTSCGSLGNIHHPGGG 652
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 567 QIINKKLDLSNVQSKGSDKNIKHVPGGGQVQIVYKPVDSLKVTSCGSLGNIHHPGGG 626
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 653 QVEVSEKLDKDRVQSKIGSLDNITHVPGGNGKTIETHKLTFRENAKAKTDHGAIVYK 712
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 627 QVEVSEKLDKDRVQSKIGSLDNITHVPGGNGKTIETHKLTFRENAKAKTDHGAIVYK 686
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 713 SPVVGDTSPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 687 SPVVGDTSPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 732
| | | | | : | | | | | | | | | : | | | | | | | | | |

```

RESULT 4

```

TAU_MACMU
ID TAU_MACMU STANDARD; PRT; 458 AA.
AC P57786;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
DE (Paired helical filament-tau) (PHF-tau).
GN MAPT OR TAU.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D).
RC TISSUE=Brain;
RX MEDLINE=97012131; PubMed=8858947;
RA Nelson P.T., Stefansson K., Gulcher J., Saper C.B.;
RT "Molecular evolution of tau protein: implications for Alzheimer's
RL J. Neurochem. 67:1622-1632(1996).
CC -1- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
CC -1- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; TAU A (SHOWN HERE),
CC TAU-B, TAU-C AND TAU-D; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY
CC DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE OF UP TO 4
CC EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS THE ADDITIONAL TAU/MAP
CC REPEAT. THE SEQUENCE SHOWN HERE IS THAT OF THE COMPLETE ISOFORM

```

```

CC NOT YET CHARACTERIZED.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
CC -1- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
CC CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
CC -1- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
CC CK5, GSK3, MARK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
CC MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC InterPro: IPR002955; Tau_protein.
CC InterPro: IPR001084; Tubulin_Tau.
CC Pfam: PF00418; tubulin-binding; 4.
CC PRINTS: PRO1261; TAUPROTEIN.
CC PROSITE: PS00229; TAU_MAP; 4.
CC Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
CC Phosphorylation.
CC INIT_MET 0 0 BY SIMILARITY.
CC REPEAT 261 291 TAU/MAP MOTIF 1.
CC REPEAT 292 322 TAU/MAP MOTIF 2.
CC REPEAT 323 353 TAU/MAP MOTIF 3.
CC REPEAT 354 385 TAU/MAP MOTIF 4.
CC MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
CC DISULFID 308 339 BY SIMILARITY.
CC VARSPLIC 44 101 MISSING (IN ISOFORM TAU-B AND ISOFORM
CC TAU-C).
CC VARSPLIC 185 202 MISSING (IN ISOFORM TAU-C AND ISOFORM
CC TAU-D).
CC VARSPLIC 292 322 MISSING (IN ISOFORM TAU-B).
CC SEQUENCE 458 AA; 47841 MW; 68BE63DBA30665A8 CRC64;

```

Query Match 51.3%; Score 2032.5; DB 1; Length 458;

Best Local Similarity 55.6%; Pred. No. 2.5e-71;

Matches 431; Conservative 2; Mismatches 7; Indels 335; Gaps 3;

```

QY 2 AEPROEFVEMDHAGTYGLGDRKQGGYTMHQDEGDTDAGLKESPLQTPTEDESEPPGS 61
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 1 AEPROEFVEMDHAGTYGLGDRKQGGYTMHQDEGDTDAGLKESPLQTPTEDESEELGS 60
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 62 ETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTPEIGTTAEAGIGDTPSLDEAAGH 121
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 61 ETSDAKSTPTAEDVTAPLVDERAAGEQAAAPHTPEIGTTAEAGIGDTPSLDEAAGH 120
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 122 VTQEPESGKVYQEGFLREPGLPGLSHQLMSGMPCAPLLPEGPREATRQPSGTGPEDETEGG 181
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 121 VTQ----- 123
QY 182 RHAPELLKHQLLDLHQEGPLPKGAGGKERPGSKEEVEDRDVDESSPDSPPSKASPAQ 241
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 124 ----- 123
QY 242 DGRPPQTAAREATSIPTFPAGEAIPLPVDFLSKVSTEIPASEPDGPSVGRAKQDAPLEF 301
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 124 ----- 123
QY 302 TFHVEITPNVQEAHSEHLGRAAFPGAPGEPGPEARGPSIGEDTKADLPPESEKQPA 361
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 124 ----- 123
QY 362 APRGKPVSRVPLKARMSKSDGTGSDKKAKTSTRSAKTLKNRCLSPKLTTPGSSD 421
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 124 -----ARMYSKSDGTGSDDKKA----- 141
QY 422 PLIQSPSPAVCEPPEPPSPKPHSVTSRTGSSCAKEMKLGADGKTKIATPRGAAPPGKG 481
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 142 -----KGADGKTKIATPRGAAPPGKG 163
QY 482 QANATRIAPKTPPAPKTPPSS-----GEPKSGDRSGYSSGSPGTP 523
| | | | | : | | | | | | | | | : | | | | | | | | | |
Db 164 QANATRIAPKTPPAPKTPPSSATKQVQRKPPAPETSERGEPKSGDRSGYSSGSPGTP 223
| | | | | : | | | | | | | | | : | | | | | | | | | |
QY 524 GSRSTRPSLPTREPKKVAVVTRPPKSPSAKSLQATAPVMPDLKNVSKIGSTENL 583
| | | | | : | | | | | | | | | : | | | | | | | | | |

```


Db 224 GSRSTPPLTPPAREPKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVSKSIGSTENL 283
 QY 584 KHQGGGKQVQIINKLDLNSVQSKCGSKDNKHVPGGSGVQIYKVPDLSKVTSCGSLG 643
 Db 284 KHQGGGKQVQIINKLDLNSVQSKCGSKDNKHVPGGSGVQIYKVPDLSKVTSCGSLG 343
 QY 644 NIHHKPGGGGVVEKSEKLFKDRVQSKIGSLDNIHVPGGGKNIETHKLTFRENAKAKT 703
 Db 344 NIHHKPGGGGVVEKSEKLFKDRVQSKIGSLDNIHVPGGGKNIETHKLTFRENAKAKT 403
 QY 704 DHGAEIYKSPVVGDTSPRHLSNVSGTSDIMVDSQPLATLADEVASLAKQGL 758
 Db 404 DHGAEIYKSPVVGDTSPRHLSNVSGTSDIMVDSQPLATLADEVASLAKQGL 458

RESULT 5
 TAU_BOVIN
 ID TAU_BOVIN STANDARD: PRT: 447 AA.
 AC P29172; P29173; Q28185; Q28186; Q28187; Q28188; Q28189; Q28190;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
 DE (Paired helical filament-tau) (PHF-tau).
 GN MAPT OR TAU.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS TAU-A; TAU-B; TAU-G AND TAU-H).
 RC TISSUE=Brain;
 RX MEDLINE=89261765; PubMed=2498649;
 RA Himmler A., Drechsel D., Kirschner M.W., Martin D.W. Jr.;
 RT "Tau consists of a set of proteins with repeated C-terminal
 RT microtubule-binding domains and variable N-terminal domains.";
 RL Mol. Cell. Biol. 9:1381-1388(1989).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS TAU-A TO TAU-F AND TAU-I TO TAU-T).
 RC TISSUE=Brain;
 RX MEDLINE=89261766; PubMed=2498650;
 RA Himmler A.;
 RT "Structure of the bovine tau gene: alternatively spliced transcripts
 RT generate a protein family.";
 RL Mol. Cell. Biol. 9:1389-1396(1989).
 RN [3]
 RP GLYCOSYLATION
 RX MEDLINE=97067111; PubMed=8910513;
 RA Arnold C.S., Johnson G.V.W., Cole R.N., Dong D.L.-Y., Lee M.,
 RA Hart G.W.;
 RT "The microtubule-associated protein tau is extensively modified with
 RT O-linked N-acetylglucosamine.";
 RL J. Biol. Chem. 271:28741-28744(1996).
 CC -|- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
 CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH AXONAL POLARITY IS
 CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
 CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
 CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
 CC -|- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS. IN THE
 CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
 CC -|- ALTERNATIVE PRODUCTS: AT LEAST 20 ISOFORMS; TAU-A/PBT43112 (SHOWN
 CC HERE), TAU-B/PBT43-12, TAU-C, TAU-D, TAU-E, TAU-F, TAU-G/PBT4,
 CC TAU-H/PBT7, TAU-I, TAU-J, TAU-K, TAU-L, TAU-M, TAU-N, TAU-O, TAU-
 CC P, TAU-Q, TAU-R, TAU-S AND TAU-T; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THEY DIFFER FROM EACH OTHER BY THE PRESENCE OR ABSENCE
 CC OF UP TO 6 OF THE 14 EXONS. ONE OF THESE OPTIONAL EXONS CONTAINS
 CC THE ADDITIONAL TAU/MAP REPEAT. TAU-A CDNA HAS BEEN CONSTRUCTED
 CC FROM TWO OVERLAPPING CDNAS BY THE AUTHORS OF REF.1. TAU-G AND TAU-

H SEQUENCES BEGIN WITH EXON 6 OR A PART OF IT (EXON 6 IS MISSING
 IN ISOFORMS THAT BEGIN WITH EXON 1). 3 DIFFERENT C-TERMINI ARE
 OBTAINED EITHER BY THE RETENTION OR THE SPLICING OF INTRON 13/14
 (2 DIFFERENT 5' SPLICE DONORS).
 -|- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
 -|- INDUCTION: DURING NEURITE OUTGROWTH.
 -|- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
 CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
 -|- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
 CK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
 MITOSIS). AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
 MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
 -|- PTM: O-LINKED N-ACETYLGLUCOSAMINATION AT MORE THAN 4 SITES PER
 PROTEIN. SITE-SPECIFIC OR STOICHIOMETRIC CHANGES IN GLYCOSYLATION
 MAY MODULATE TAU FUNCTION AND ALSO PLAY A ROLE IN PHF'S FORMATION.
 -|- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; L34953; AAA51609.1; -
 EMBL; L34940; AAA51609.1; JOINED.
 EMBL; L34941; AAA51609.1; JOINED.
 EMBL; L34942; AAA51609.1; JOINED.
 EMBL; L34943; AAA51609.1; JOINED.
 EMBL; L34944; AAA51609.1; JOINED.
 EMBL; L34946; AAA51609.1; JOINED.
 EMBL; L34947; AAA51609.1; JOINED.
 EMBL; L34948; AAA51609.1; JOINED.
 EMBL; L34949; AAA51609.1; JOINED.
 EMBL; L34950; AAA51609.1; JOINED.
 EMBL; L34951; AAA51609.1; JOINED.
 EMBL; L34952; AAA51609.1; JOINED.
 EMBL; M26157; AAA30770.1; -
 EMBL; M26158; AAA30771.1; -
 EMBL; M26178; AAA51601.1; ALT_SEQ.
 EMBL; L34940; AAA51601.1; JOINED.
 EMBL; L34941; AAA51601.1; JOINED.
 EMBL; L34942; AAA51601.1; JOINED.
 EMBL; L34943; AAA51601.1; JOINED.
 EMBL; L34944; AAA51601.1; JOINED.
 EMBL; L34946; AAA51601.1; JOINED.
 EMBL; L34947; AAA51601.1; JOINED.
 EMBL; L34948; AAA51601.1; JOINED.
 EMBL; L34949; AAA51601.1; JOINED.
 EMBL; L34950; AAA51601.1; JOINED.
 EMBL; L34951; AAA51601.1; JOINED.
 EMBL; L34952; AAA51601.1; JOINED.
 EMBL; L34953; AAA51601.1; JOINED.
 EMBL; L34954; AAA51601.1; JOINED.
 EMBL; L34955; AAA51601.1; JOINED.
 EMBL; L34956; AAA51601.1; JOINED.
 EMBL; L34957; AAA51601.1; JOINED.
 EMBL; L34958; AAA51601.1; JOINED.
 EMBL; L34959; AAA51601.1; JOINED.
 EMBL; L34960; AAA51601.1; JOINED.
 EMBL; L34961; AAA51601.1; JOINED.
 EMBL; L34962; AAA51601.1; JOINED.
 EMBL; L34963; AAA51601.1; JOINED.
 EMBL; L34964; AAA51601.1; JOINED.
 EMBL; L34965; AAA51601.1; JOINED.
 EMBL; L34966; AAA51601.1; JOINED.
 EMBL; L34967; AAA51601.1; JOINED.
 EMBL; L34968; AAA51601.1; JOINED.
 EMBL; L34969; AAA51601.1; JOINED.
 EMBL; L34970; AAA51601.1; JOINED.
 EMBL; L34971; AAA51601.1; JOINED.
 EMBL; L34972; AAA51601.1; JOINED.
 EMBL; L34973; AAA51601.1; JOINED.
 EMBL; L34974; AAA51601.1; JOINED.
 EMBL; L34975; AAA51601.1; JOINED.
 EMBL; L34976; AAA51601.1; JOINED.
 EMBL; L34977; AAA51601.1; JOINED.
 EMBL; L34978; AAA51601.1; JOINED.
 EMBL; L34979; AAA51601.1; JOINED.
 EMBL; L34980; AAA51601.1; JOINED.
 EMBL; L34981; AAA51601.1; JOINED.
 EMBL; L34982; AAA51601.1; JOINED.
 EMBL; L34983; AAA51601.1; JOINED.
 EMBL; L34984; AAA51601.1; JOINED.
 EMBL; L34985; AAA51601.1; JOINED.
 EMBL; L34986; AAA51601.1; JOINED.
 EMBL; L34987; AAA51601.1; JOINED.
 EMBL; L34988; AAA51601.1; JOINED.
 EMBL; L34989; AAA51601.1; JOINED.
 EMBL; L34990; AAA51601.1; JOINED.
 EMBL; L34991; AAA51601.1; JOINED.
 EMBL; L34992; AAA51601.1; JOINED.
 EMBL; L34993; AAA51601.1; JOINED.
 EMBL; L34994; AAA51601.1; JOINED.
 EMBL; L34995; AAA51601.1; JOINED.
 EMBL; L34996; AAA51601.1; JOINED.
 EMBL; L34997; AAA51601.1; JOINED.
 EMBL; L34998; AAA51601.1; JOINED.
 EMBL; L34999; AAA51601.1; JOINED.
 EMBL; L35000; AAA51601.1; JOINED.
 EMBL; L35001; AAA51601.1; JOINED.
 EMBL; L35002; AAA51601.1; JOINED.
 EMBL; L35003; AAA51601.1; JOINED.
 EMBL; L35004; AAA51601.1; JOINED.
 EMBL; L35005; AAA51601.1; JOINED.
 EMBL; L35006; AAA51601.1; JOINED.
 EMBL; L35007; AAA51601.1; JOINED.
 EMBL; L35008; AAA51601.1; JOINED.
 EMBL; L35009; AAA51601.1; JOINED.
 EMBL; L35010; AAA51601.1; JOINED.
 EMBL; L35011; AAA51601.1; JOINED.
 EMBL; L35012; AAA51601.1; JOINED.
 EMBL; L35013; AAA51601.1; JOINED.
 EMBL; L35014; AAA51601.1; JOINED.
 EMBL; L35015; AAA51601.1; JOINED.
 EMBL; L35016; AAA51601.1; JOINED.
 EMBL; L35017; AAA51601.1; JOINED.
 EMBL; L35018; AAA51601.1; JOINED.
 EMBL; L35019; AAA51601.1; JOINED.
 EMBL; L35020; AAA51601.1; JOINED.
 EMBL; L35021; AAA51601.1; JOINED.
 EMBL; L35022; AAA51601.1; JOINED.
 EMBL; L35023; AAA51601.1; JOINED.
 EMBL; L35024; AAA51601.1; JOINED.
 EMBL; L35025; AAA51601.1; JOINED.
 EMBL; L35026; AAA51601.1; JOINED.
 EMBL; L35027; AAA51601.1; JOINED.
 EMBL; L35028; AAA51601.1; JOINED.
 EMBL; L35029; AAA51601.1; JOINED.
 EMBL; L35030; AAA51601.1; JOINED.
 EMBL; L35031; AAA51601.1; JOINED.
 EMBL; L35032; AAA51601.1; JOINED.
 EMBL; L35033; AAA51601.1; JOINED.
 EMBL; L35034; AAA51601.1; JOINED.
 EMBL; L35035; AAA51601.1; JOINED.
 EMBL; L35036; AAA51601.1; JOINED.
 EMBL; L35037; AAA51601.1; JOINED.
 EMBL; L35038; AAA51601.1; JOINED.
 EMBL; L35039; AAA51601.1; JOINED.
 EMBL; L35040; AAA51601.1; JOINED.
 EMBL; L35041; AAA51601.1; JOINED.
 EMBL; L35042; AAA51601.1; JOINED.
 EMBL; L35043; AAA51601.1; JOINED.
 EMBL; L35044; AAA51601.1; JOINED.
 EMBL; L35045; AAA51601.1; JOINED.
 EMBL; L35046; AAA51601.1; JOINED.
 EMBL; L35047; AAA51601.1; JOINED.
 EMBL; L35048; AAA51601.1; JOINED.
 EMBL; L35049; AAA51601.1; JOINED.
 EMBL; L35050; AAA51601.1; JOINED.
 EMBL; L35051; AAA51601.1; JOINED.
 EMBL; L35052; AAA51601.1; JOINED.
 EMBL; L35053; AAA51601.1; JOINED.
 EMBL; L35054; AAA51601.1; JOINED.
 EMBL; L35055; AAA51601.1; JOINED.
 EMBL; L35056; AAA51601.1; JOINED.
 EMBL; L35057; AAA51601.1; JOINED.
 EMBL; L35058; AAA51601.1; JOINED.
 EMBL; L35059; AAA51601.1; JOINED.
 EMBL; L35060; AAA51601.1; JOINED.
 EMBL; L35061; AAA51601.1; JOINED.
 EMBL; L35062; AAA51601.1; JOINED.
 EMBL; L35063; AAA51601.1; JOINED.
 EMBL; L35064; AAA51601.1; JOINED.
 EMBL; L35065; AAA51601.1; JOINED.
 EMBL; L35066; AAA51601.1; JOINED.
 EMBL; L35067; AAA51601.1; JOINED.
 EMBL; L35068; AAA51601.1; JOINED.
 EMBL; L35069; AAA51601.1; JOINED.
 EMBL; L35070; AAA51601.1; JOINED.
 EMBL; L35071; AAA51601.1; JOINED.
 EMBL; L35072; AAA51601.1; JOINED.
 EMBL; L35073; AAA51601.1; JOINED.
 EMBL; L35074; AAA51601.1; JOINED.
 EMBL; L35075; AAA51601.1; JOINED.
 EMBL; L35076; AAA51601.1; JOINED.
 EMBL; L35077; AAA51601.1; JOINED.
 EMBL; L35078; AAA51601.1; JOINED.
 EMBL; L35079; AAA51601.1; JOINED.
 EMBL; L35080; AAA51601.1; JOINED.
 EMBL; L35081; AAA51601.1; JOINED.
 EMBL; L35082; AAA51601.1; JOINED.
 EMBL; L35083; AAA51601.1; JOINED.
 EMBL; L35084; AAA51601.1; JOINED.
 EMBL; L35085; AAA51601.1; JOINED.
 EMBL; L35086; AAA51601.1; JOINED.
 EMBL; L35087; AAA51601.1; JOINED.
 EMBL; L35088; AAA51601.1; JOINED.
 EMBL; L35089; AAA51601.1; JOINED.
 EMBL; L35090; AAA51601.1; JOINED.
 EMBL; L35091; AAA51601.1; JOINED.
 EMBL; L35092; AAA51601.1; JOINED.
 EMBL; L35093; AAA51601.1; JOINED.
 EMBL; L35094; AAA51601.1; JOINED.
 EMBL; L35095; AAA51601.1; JOINED.
 EMBL; L35096; AAA51601.1; JOINED.
 EMBL; L35097; AAA51601.1; JOINED.
 EMBL; L35098; AAA51601.1; JOINED.
 EMBL; L35099; AAA51601.1; JOINED.
 EMBL; L35100; AAA51601.1; JOINED.
 EMBL; L35101; AAA51601.1; JOINED.
 EMBL; L35102; AAA51601.1; JOINED.
 EMBL; L35103; AAA51601.1; JOINED.
 EMBL; L35104; AAA51601.1; JOINED.
 EMBL; L35105; AAA51601.1; JOINED.
 EMBL; L35106; AAA51601.1; JOINED.
 EMBL; L35107; AAA51601.1; JOINED.
 EMBL; L35108; AAA51601.1; JOINED.
 EMBL; L35109; AAA51601.1; JOINED.
 EMBL; L35110; AAA51601.1; JOINED.
 EMBL; L35111; AAA51601.1; JOINED.
 EMBL; L35112; AAA51601.1; JOINED.
 EMBL; L35113; AAA51601.1; JOINED.
 EMBL; L35114; AAA51601.1; JOINED.
 EMBL; L35115; AAA51601.1; JOINED.
 EMBL; L35116; AAA51601.1; JOINED.
 EMBL; L35117; AAA51601.1; JOINED.
 EMBL; L35118; AAA51601.1; JOINED.
 EMBL; L35119; AAA51601.1; JOINED.
 EMBL; L35120; AAA51601.1; JOINED.
 EMBL; L35121; AAA51601.1; JOINED.
 EMBL; L35122; AAA51601.1; JOINED.
 EMBL; L35123; AAA51601.1; JOINED.
 EMBL; L35124; AAA51601.1; JOINED.
 EMBL; L35125; AAA51601.1; JOINED.
 EMBL; L35126; AAA51601.1; JOINED.
 EMBL; L35127; AAA51601.1; JOINED.
 EMBL; L35128; AAA51601.1; JOINED.
 EMBL; L35129; AAA51601.1; JOINED.
 EMBL; L35130; AAA51601.1; JOINED.
 EMBL; L35131; AAA51601.1; JOINED.
 EMBL; L35132; AAA51601.1; JOINED.
 EMBL; L35133; AAA51601.1; JOINED.
 EMBL; L35134; AAA51601.1; JOINED.
 EMBL; L35135; AAA51601.1; JOINED.
 EMBL; L35136; AAA51601.1; JOINED.
 EMBL; L35137; AAA51601.1; JOINED.
 EMBL; L35138; AAA51601.1; JOINED.
 EMBL; L35139; AAA51601.1; JOINED.
 EMBL; L35140; AAA51601.1; JOINED.
 EMBL; L35141; AAA51601.1; JOINED.
 EMBL; L35142; AAA51601.1; JOINED.
 EMBL; L35143; AAA51601.1; JOINED.
 EMBL; L35144; AAA51601.1; JOINED.
 EMBL; L35145; AAA51601.1; JOINED.
 EMBL; L35146; AAA51601.1; JOINED.
 EMBL; L35147; AAA51601.1; JOINED.
 EMBL; L35148; AAA51601.1; JOINED.
 EMBL; L35149; AAA51601.1; JOINED.
 EMBL; L35150; AAA51601.1; JOINED.
 EMBL; L35151; AAA51601.1; JOINED.
 EMBL; L35152; AAA51601.1; JOINED.
 EMBL; L35153; AAA51601.1; JOINED.
 EMBL; L35154; AAA51601.1; JOINED.
 EMBL; L35155; AAA51601.1; JOINED.
 EMBL; L35156; AAA51601.1; JOINED.
 EMBL; L35157; AAA51601.1; JOINED.
 EMBL; L35158; AAA51601.1; JOINED.
 EMBL; L35159; AAA51601.1; JOINED.
 EMBL; L35160; AAA51601.1; JOINED.
 EMBL; L35161; AAA51601.1; JOINED.
 EMBL; L35162; AAA51601.1; JOINED.
 EMBL; L35163; AAA51601.1; JOINED.
 EMBL; L35164; AAA51601.1; JOINED.
 EMBL; L35165; AAA51601.1; JOINED.
 EMBL; L35166; AAA51601.1; JOINED.
 EMBL; L35167; AAA51601.1; JOINED.
 EMBL; L35168; AAA51601.1; JOINED.
 EMBL; L35169; AAA51601.1; JOINED.
 EMBL; L35170; AAA51601.1; JOINED.
 EMBL; L35171; AAA51601.1; JOINED.
 EMBL; L35172; AAA51601.1; JOINED.
 EMBL; L35173; AAA51601.1; JOINED.
 EMBL; L35174; AAA51601.1; JOINED.
 EMBL; L35175; AAA51601.1; JOINED.
 EMBL; L35176; AAA51601.1; JOINED.
 EMBL; L35177; AAA51601.1; JOINED.
 EMBL; L35178; AAA51601.1; JOINED.
 EMBL; L35179; AAA51601.1; JOINED.
 EMBL; L35180; AAA51601.1; JOINED.
 EMBL; L35181; AAA51601.1; JOINED.
 EMBL; L35182; AAA51601.1; JOINED.
 EMBL; L35183; AAA51601.1; JOINED.
 EMBL; L35184; AAA51601.1; JOINED.
 EMBL; L35185; AAA51601.1; JOINED.
 EMBL; L35186; AAA51601.1; JOINED.
 EMBL; L35187; AAA51601.1; JOINED.
 EMBL; L35188; AAA51601.1; JOINED.
 EMBL; L35189; AAA51601.1; JOINED.
 EMBL; L35190; AAA51601.1; JOINED.
 EMBL; L35191; AAA51601.1; JOINED.
 EMBL; L35192; AAA51601.1; JOINED.
 EMBL; L35193; AAA51601.1; JOINED.
 EMBL; L35194; AAA51601.1; JOINED.
 EMBL; L35195; AAA51601.1; JOINED.
 EMBL; L35196; AAA51601.1; JOINED.
 EMBL; L35197; AAA51601.1; JOINED.
 EMBL; L35198; AAA51601.1; JOINED.
 EMBL; L35199; AAA51601.1; JOINED.
 EMBL; L35200; AAA51601.1; JOINED.
 EMBL; L35201; AAA51601.1; JOINED.
 EMBL; L35202; AAA51601.1; JOINED.
 EMBL; L35203; AAA51601.1; JOINED.
 EMBL; L35204; AAA51601.1; JOINED.
 EMBL; L35205; AAA51601.1; JOINED.
 EMBL; L35206; AAA51601.1; JOINED.
 EMBL; L35207; AAA51601.1; JOINED.
 EMBL; L35208; AAA51601.1; JOINED.
 EMBL; L35209; AAA51601.1; JOINED.
 EMBL; L35210; AAA51601.1; JOINED.
 EMBL; L35211; AAA51601.1; JOINED.
 EMBL; L35212; AAA51601.1; JOINED.
 EMBL; L35213; AAA51601.1; JOINED.
 EMBL; L35214; AAA51601.1; JOINED.
 EMBL; L35215; AAA51601.1; JOINED.
 EMBL; L35216; AAA51601.1; JOINED.
 EMBL; L35217; AAA51601.1; JOINED.
 EMBL; L35218; AAA51601.1; JOINED.
 EMBL; L35219; AAA51601.1; JOINED.
 EMBL; L35220; AAA51601.1; JOINED.
 EMBL; L35221; AAA51601.1; JOINED.
 EMBL; L35222; AAA51601.1; JOINED.
 EMBL; L35223; AAA51601.1; JOINED.
 EMBL; L35224; AAA51601.1; JOINED.
 EMBL; L35225; AAA51601.1; JOINED.
 EMBL; L35226; AAA51601.1; JOINED.
 EMBL; L35227; AAA51601.1; JOINED.
 EMBL; L35228; AAA51601.1; JOINED.
 EMBL; L35229; AAA51601.1; JOINED.
 EMBL; L35230; AAA51601.1; JOINED.
 EMBL; L35231; AAA51601.1; JOINED.
 EMBL; L35232; AAA51601.1; JOINED.
 EMBL; L35233; AAA51601.1; JOINED.
 EMBL; L35234; AAA51601.1; JOINED.
 EMBL; L35235; AAA51601.1; JOINED.
 EMBL; L35236; AAA51601.1; JOINED.
 EMBL; L35237; AAA51601.1; JOINED.
 EMBL; L35238; AAA51601.1; JOINED.
 EMBL; L35239; AAA51601.1; JOINED.
 EMBL; L35240; AAA51601.1; JOINED.
 EMBL; L35241; AAA51601.1; JOINED.
 EMBL; L35242; AAA51601.1; JOINED.
 EMBL; L35243; AAA51601.1; JOINED.
 EMBL; L35244; AAA51601.1; JOINED.
 EMBL; L35245; AAA51601.1; JOINED.
 EMBL; L35246; AAA51601.1; JOINED.
 EMBL; L35247; AAA51601.1; JOINED.
 EMBL; L35248; AAA51601.1; JOINED.
 EMBL; L35249; AAA51601.1; JOINED.
 EMBL; L35250; AAA51601.1; JOINED.
 EMBL; L35251; AAA51601.1; JOINED.
 EMBL; L35252; AAA51601.1; JOINED.
 EMBL; L35253; AAA51601.1; JOINED.
 EMBL; L35254; AAA51601.1; JOINED.
 EMBL; L35255; AAA51601.1; JOINED.
 EMBL; L35256; AAA51601.1; JOINED.
 EMBL; L35257; AAA51601.1; JOINED.
 EMBL; L35258; AAA51601.1; JOINED.
 EMBL; L35259; AAA51601.1; JOINED.
 EMBL; L35260; AAA51601.1; JOINED.
 EMBL; L35261; AAA51601.1; JOINED.
 EMBL; L35262; AAA51601.1; JOINED.
 EMBL; L35263; AAA51601.1; JOINED.
 EMBL; L35264; AAA51601.1; JOINED.
 EMBL; L35265; AAA51601.1; JOINED.
 EMBL; L35266; AAA51601.1; JOINED.
 EMBL; L35267; AAA51601.1; JOINED.
 EMBL; L35268; AAA51601.1; JOINED.
 EMBL; L35269; AAA51601.1; JOINED.
 EMBL; L35270; AAA51601.1; JOINED.
 EMBL; L35271; AAA51601.1; JOINED.
 EMBL; L35272; AAA51601.1; JOINED.
 EMBL; L35273; AAA51601.1; JOINED.
 EMBL; L35274; AAA51601.1; JOINED.
 EMBL; L35275; AAA51601.1; JOINED.
 EMBL; L35276; AAA51601.1; JOINED.
 EMBL; L35277; AAA51601.1; JOINED.
 EMBL; L35278; AAA51601.1; JOINED.
 EMBL; L35279; AAA51601.1; JOINED.
 EMBL; L35280; AAA51601.1; JOINED.
 EMBL; L35281; AAA51601.1; JOINED.
 EMBL; L35282; AAA51601.1; JOINED.
 EMBL; L35283; AAA51601.1; JOINED.
 EMBL; L35284; AAA51601.1; JOINED.
 EMBL; L35285; AAA51601.1; JOINED.
 EMBL; L35286; AAA51601.1; JOINED.
 EMBL; L35287; AAA51601.1; JOINED.
 EMBL; L35288; AAA51601.1; JOINED.
 EMBL; L35289; AAA51601.1; JOINED.
 EMBL; L35290; AAA51601.1; JOINED.
 EMBL; L35291; AAA51601.1; JOINED.
 EMBL; L35292; AAA51601

RC TISSUE=Frontal cortex;
 RA Wang X.L., Wang J., Schultz C., Hubbard G.B.;
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
 CC -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
 CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
 CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
 CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS (BY
 CC SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
 CC -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN.
 CC -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,
 CC CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
 CC MITOSIS), AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
 CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF281310; AAF97596.1; -;
 DR InterPro: IPR002955; Tau_protein.
 DR InterPro: IPR001084; Tubulin_Tau.
 DR Pfam: PF00418; tubulin-binding; 4.
 DR PRINTS: PR01261; TAUPROTEIN.
 DR PROSITE: PS00229; TAU_MAP; 4.
 KW Microtubules; Cytoskeleton; Repeat; Acetylation; Phosphorylation.
 FT REPEAT 0 0 BY SIMILARITY.
 FT INIT_MET 185 215 TAU/MAP MOTIF 1.
 FT REPEAT 216 246 TAU/MAP MOTIF 2.
 FT REPEAT 247 277 TAU/MAP MOTIF 3.
 FT REPEAT 278 309 TAU/MAP MOTIF 4.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 FT DISULFID 232 263 BY SIMILARITY.
 FT SEQUENCE 382 AA; 39879 MW; D2D15A53AA00E8E7 CRC64;
 SQ
 Query Match 44.0%; Score 1744.5; DB 1; Length 392;
 Best Local Similarity 49.9%; Pred. No. 1.9e-60;
 Matches 378; Conservative 1; Mismatches 3; Indels 375; Gaps 3;
 QY 2 AEPQEFVMEHDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQTPTEDGSEEPGS 61
 DB 1 AEPQEFVMEHDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQTPTEDGSEEPGS 61
 QY 62 ETSDAKSTPAEDVTAPLVDEGAPGQAAAPHTPEIGTTAAEAGIGDTPSLEDAAGH 121
 DB 44 ETSDAKSTPAEDVTAPLVDEGAPGQAAAPHTPEIGTTAAEAGIGDTPSLEDAAGH 121
 QY 122 VTQEPESGKVGQEGFLREPPGGLSHQLMSGMGAPLLPEGPREATRPGSGTGEDTEGG 181
 DB 63 VTQEPESGKVGQEGFLREPPGGLSHQLMSGMGAPLLPEGPREATRPGSGTGEDTEGG 181
 QY 182 RHAPELLKQLGLDLHQEGPPLKAGGKRPCKEEDVEDRDVDESSPDSPSKASPAQ 241
 DB 66 RHAPELLKQLGLDLHQEGPPLKAGGKRPCKEEDVEDRDVDESSPDSPSKASPAQ 241
 QY 242 DGRPPQTAREATSTPGPPAEGATPLPVDLSKYSTPEIPASEPDGSPVGRAGQADPLEF 301
 DB 66 DGRPPQTAREATSTPGPPAEGATPLPVDLSKYSTPEIPASEPDGSPVGRAGQADPLEF 301
 QY 302 TFHVEITPNVQEAHSEHLGRAAFPAGCGEPARGPSLGEDTKEADLPSEKQPA 361
 DB 66 TFHVEITPNVQEAHSEHLGRAAFPAGCGEPARGPSLGEDTKEADLPSEKQPA 361

QY 362 APRGKPVSRVLPOLKARMVSKSKDGTGSDKKAKTSTRSSAKTLKRNPCLSPLPTPGSSD 421
 DB 66 APRGKPVSRVLPOLKARMVSKSKDGTGSDKKAKTSTRSSAKTLKRNPCLSPLPTPGSSD 421
 QY 422 PLIQSSPAVCPPEPSSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGKG 481
 DB 84 PLIQSSPAVCPPEPSSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGKG 481
 QY 482 QANATRIPAKTPPAPKTPPSGEPKSGDRSGYSPGSPGTPGSRSTRPSLPTPTREP 541
 DB 106 QANATRIPAKTPPAPKTPPSGEPKSGDRSGYSPGSPGTPGSRSTRPSLPTPTREP 541
 QY 542 KVAVVTRTPPKSPSSAKSLQATAPVMPDLKNVSKSGISTENLKHQPGGQGVKSEKL 601
 DB 166 KVAVVTRTPPKSPSSAKSLQATAPVMPDLKNVSKSGISTENLKHQPGGQGVKSEKL 601
 QY 602 SNVQSKGSKDNKIHVPKGGSVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSEKL 661
 DB 226 SNVQSKGSKDNKIHVPKGGSVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSEKL 661
 QY 662 DFKDRVQSKIGSLDNITHVPGGGNKKIETHKLTFRENAKAKTDHGAETVYKSPVVGDT 721
 DB 286 DFKDRVQSKIGSLDNITHVPGGGNKKIETHKLTFRENAKAKTDHGAETVYKSPVVGDT 721
 QY 722 PRHLSNVSTGSDIDWVSPQIATLADEVSAKOG 758
 DB 346 PRHLSNVSTGSDIDWVSPQIATLADEVSAKOG 758
 RESULT 7
 TAU_CAPHI
 ID TAU_CAPHI STANDARD; PRT; 402 AA.
 AC 002828;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Microtubule-associated protein tau (Neurofibrillary tangle protein)
 DE (Paired helical filament-tau) (PHF-tau).
 GN MAPT OR TAU.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9225;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
 RC TISSUE=Brain cortex;
 RX MEDLINE=97012131; PubMed=8858947;
 RA Nelson P.T., Stefansson K., Gulcher J., Saper C.B.;
 RT "Molecular evolution of tau protein: Implications for Alzheimer's
 RT disease.";
 RL J. Neurochem. 67:1622-1632(1996).
 CC -!- FUNCTION: PROMOTES MICROTUBULE ASSEMBLY AND STABILITY, AND MIGHT
 CC BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL
 CC POLARITY. THE C-TERMINUS BINDS AXONAL MICROTUBULES WHILE THE N-
 CC TERMINUS BINDS NEURAL PLASMA MEMBRANE COMPONENTS, SUGGESTING THAT
 CC TAU FUNCTIONS AS A LINKER PROTEIN BETWEEN BOTH. AXONAL POLARITY IS
 CC PREDETERMINED BY TAU LOCALIZATION (IN THE NEURONAL CELL) IN THE
 CC DOMAIN OF THE CELL BODY DEFINED BY THE CENTROSOME. THE SHORT
 CC ISOFORMS ALLOW PLASTICITY OF THE CYTOSKELETON WHEREAS THE LONGER
 CC ISOFORMS MAY PREFERENTIALLY PLAY A ROLE IN ITS STABILIZATION.
 CC -!- SUBCELLULAR LOCATION: MOSTLY FOUND IN THE AXONS OF NEURONS, IN THE
 CC CYTOSOL AND IN ASSOCIATION WITH PLASMA MEMBRANE COMPONENTS.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; TAU-A (SHOWN HERE) AND
 CC TAU-B; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER FROM EACH
 CC OTHER BY THE PRESENCE OR ABSENCE OF TWO EXONS, ONE OF THESE
 CC OPTIONAL EXONS CONTAINS THE ADDITIONAL TAU/MAP REPEAT.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONS.
 CC -!- DOMAIN: THE TAU/MAP REPEAT BINDS TO TUBULIN. TYPE I ISOFORMS
 CC CONTAIN 3 REPEATS WHILE TYPE II ISOFORMS CONTAIN 4 REPEATS.
 CC -!- PTM: PHOSPHORYLATION AT VARIOUS SERINE AND THREONINE RESIDUES IN
 CC S-P OR T-P MOTIFS BY PROLINE-DIRECTED PROTEIN KINASES (PDPK: CDC2,

CC CDK5, GSK3, MAPK) (A FEW SITES PER PROTEIN IN INTERPHASE, MORE IN
CC MITOSIS) AND AT SERINE RESIDUES IN K-X-G-S MOTIFS BY
CC MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK) (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S83347; AAB50785.1; -;
DR InterPro: IPR002955; Tau_protein.
DR InterPro: IPR001084; Tubulin_tau.
DR Pfam: PF00418; tubulin-binding; 4.
DR PRINTS: PRO1261; TAUPROTEIN.
DR PROSITE: PS00229; TAU MAP; 4.
KW Microtubules; Cytoskeleton; Repeat; Alternative splicing; Acetylation;
KW Phosphorylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT REPEAT 205 235 TAU/MAP MOTIF 1.
FT REPEAT 236 266 TAU/MAP MOTIF 2.
FT REPEAT 267 297 TAU/MAP MOTIF 3.
FT REPEAT 298 329 TAU/MAP MOTIF 4.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT DISULFID 252 283 BY SIMILARITY.
FT VARSPPLIC 33 61 MISSING (IN ISOFORM TAU-B).
FT VARSPPLIC 236 266 MISSING (IN ISOFORM TAU-B).
SQ SEQUENCE 402 AA; 41716 MW; 3623B584E9F8AEF CRC64;

Query Match 43.2%; Score 1711.5; DB 1; Length 402;
Best Local Similarity 49.5%; Pred. No. 3.6e-59;
Matches 376; Conservative 8; Mismatches 16; Indels 359; Gaps 6;

QY 2 AEPQEFVEMDHAGTYGLDRKDDGGYTMHQDQEGDTAGLKESPLQPTTGGSEEPGS 61
DB 1 AEPQEFVEMDHA-----QGDYTL-QDHEGDMPEGLKESPLQPTADGGSEEPGS 49
QY 62 ETSDAKSTPTAEDVTAFLVDEGAPGQAQAAPHTPEGTTFEEAGIGTPTSLDENAGH 121
DB 50 ETSDAKSTPTAE-----AEEAGIGTSTNLEDAQAAGH 80
QY 122 VTQEPESGKVVQEGFLREPGLSHQLNSGMPGAPLLPEGPREATRQPSGTGPEDEGG 181
DB 81 VTQ----- 83
QY 182 RHAPELKHOLLGLDHOEGPLKAGGKERPGSKEEVEDRDVDSESSPODSPSKASPAQ 241
DB 84 ----- 83
QY 242 DGRPPQTAAREATISGPPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAKQDAPLEF 301
DB 84 ----- 83
QY 302 TFHVEITPNVQKEQAHSEHLGRAFFGAPGEGPEARGLSGEDTREADLPPESEKQAAA 361
DB 84 ----- 83
QY 362 APRGKPVSRVQLKARVSKSDGTGSDDKKAKTSTRSSAKTLKRNPLSPKLPTPGSSD 421
DB 84 -----ARMVSKGDTGTGDDKKA----- 101
QY 422 PLIQSPSPAVCPPEPPSPKHSVSTSGSGGAKMKLGADCK--TKIATPRGAAPQG 479
DB 102 -----KGADGKPGTKATPRGAAPQG 123
QY 480 KQANATRIPAKTPPAPKTPPSGCEPKSGDRSGSPGTPGSRSTPSPPTPTRE 539
DB 124 KQANATRIPAKTPPAPKTPPSGCEPKSGDRSGSPGTPGSRSTPSPPTPTRE 183
QY 540 PKXVAVVTRTPPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKVQIINKKL 599

DB 184 PKXVAVVTRTPPKSPSSAKSLQAPGMPDLKNVKSIGSTENLKHQPGGKVQIINKKL 243
QY 600 DLSNVQSKCGSKDNKIKHVPGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSE 659
DB 244 DLSNVQSKCGSKDNKIKHVPGGSGVQIVYKPDLSKVTSCGSLGNTHHKPGGQGVKSE 303
QY 660 KLDFKDRVQSKIGSLDNIHVPGGGNKKIETHKLTFRENAAKAKTDHGAEIVYKSPVVS 719
DB 304 KLDFKDRVQSKIGSLDNIHVPGGGNKKIETHKLTFRENAAKAKTDHGAEIVYKSPVVS 363
QY 720 TSPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
DB 364 TSPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 402

RESULT 8
MAP2_RAT
ID MAP2_RAT STANDARD; PRT; 1861 AA.
AC P15146;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2) (MAP2B) (Contains: MAP2C).
GN MAP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=90251471; PubMed=2339070;
RA Kindler S., Schwanke B., Schulz B., Garner C.C.;
RT "Complete cDNA sequence encoding rat high and low molecular weight
MAP2".
RL Nucleic Acids Res. 18:2822-2822(1990).
RN [2]
RP SEQUENCE OF 1-1694 AND 1726-1861 FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=91060576; PubMed=2174050;
RA Kindler S., Schulz B., Goedert M., Garner C.C.;
RT "Molecular structure of microtubule-associated protein 2b and 2c from
rat brain".
RL J. Biol. Chem. 265:19679-19684(1990).
RN [3]
RP SEQUENCE OF 1-151; 1515-1694 AND 1726-1861 FROM N.A.
RX MEDLINE=90221819; PubMed=2326166;
RA Doll T., Papandrikopoulou A., Matus A.;
RT "Nucleotide and amino acid sequences of embryonic rat MAP2c".
RL Nucleic Acids Res. 18:361-361(1990).
RN [4]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=89365159; PubMed=2770869;
RA Papandrikopoulou A., Doll T., Tucker R.P., Garner C.C., Matus A.;
RT "Embryonic MAP2 lacks the cross-linking sidearm sequences and
dendritic targeting signal of adult MAP2".
RL Nature 340:650-652(1989).
RN [5]
RP SEQUENCE OF 1695-1725 FROM N.A.
RX MEDLINE=94110302; PubMed=8282767;
RA Doll T., Meichner M., Riederer B.M., Honegger P., Matus A.;
RT "An isoform of microtubule-associated protein 2 (MAP2) containing
four repeats of the tubulin-binding motif".
RL J. Cell Sci. 106:633-640(1993).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF MAP2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF THE SAME GENE. MAP2C, THE LOW MOLECULAR
CC FORM OF MAP2, LACKS THE CENTRAL DOMAIN OF MAP2A/B.
CC -!- DEVELOPMENTAL STAGE: MAP2C IS EXPRESSED DURING EMBRYONIC BRAIN
CC DEVELOPMENT AND UNTIL POSTNATAL DAY 10. MAP2B IS EXPRESSED

or send an email to license@isb-sib.ch).

```

CC CC EMBL; U01828; AAA03354.1; -
DR DR EMBL; U89330; AAB48098.1; -
DR DR EMBL; U893329; AAB48097.1; -
DR DR EMBL; M25668; AAA59552.1; -
DR DR PIR; PL0024; ORHUNT.
DR DR Genew; HGNC:6839; MAP2.
DR DR MIM; 157130; -
DR DR InterPro; IPR001084; Tubulin_Tau.
DR DR Pfam; PF00418; tubulin-binding; 3.
DR DR PROSITE; PS00229; TAU_MAP; 2.
KW Microtubules; Repeat; Alternative splicing; Calmodulin-binding.
FT DOMAIN 1447 1467 CALMODULIN-BINDING (POTENTIAL).
FT REPEAT 1661 1691 TAU/MAP MOTIF.
FT REPEAT 1692 1722 TAU/MAP MOTIF.
FT REPEAT 1723 1754 TAU/MAP MOTIF.
FT VARSPLIC 152 1507 MISSING (IN ISOFORM MAP2C).
FT CONFLICT 9 9 A -> G (IN REF. 2).
FT CONFLICT 37 37 A -> G (IN REF. 2).
FT CONFLICT 108 108 A -> G (IN REF. 2).
FT CONFLICT 152 155 MISSING (IN REF. 2).
FT CONFLICT 187 187 S -> K (IN REF. 2).
FT CONFLICT 1655 1655 A -> GL (IN REF. 2).
FT CONFLICT 1736 1736 V -> A (IN REF. 2).
SQ SEQUENCE 1827 AA; 199610 MW; BAC36D0030F5F455 CRC64;

Query Match 20.6%; Score 815; DB 1; Length 1827;
Best Local Similarity 29.4%; Pred. No. 1.6e-24;
Matches 286; Conservative 122; Mismatches 274; Indels 292; Gaps 40;

QY 5 RQFEYMEHAGTYGLGDRKGDTMDQEGDTDAGLKESPLQPTEDGSEEGSETS 64
DB 926 KDFSV-DKEASHISDKR---SGLSKFEDQKKAN---DRLDTVLEK-SEEHADSK 975
QY 65 DAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPECTTAEEA--GIGDTPSLEDEAAGHV 122
DB 976 HAKKTEAGD-ETETGGLVTEQALAK-DLSIPTDASSEKAEGLSVPEI----- 1025
QY 123 TQEPESGKVVOEGF-----LREPGPGPLSHOLSMGM-----PCAPLLPEGPREA 166
DB 1026 -AEVPSKKVEGLDFAVOGLDVKISDFG-----QMASGLNIDRRATELKLEATQDM 1078
QY 167 TROPSTGTGPD-----EGGRHAPELLKHLQGLD--LHQEGPLK---GAGGK 209
DB 1079 T--PSSKAQEQADARNGVESGHMKGCTKVSETVKQKVPDLVHQEAVDKESYESGE 1136
QY 210 ERPGSKVEEDRDVDESSPDQD-----PP--SKASPAQDGRP----- 245
DB 1137 HESLTWESLKADEGKETSPESSLIQDEIAVKLSVEIPCPVASEADLATDERADVQMEF 1196
QY 246 ---PQTAAREATSIQCPAEGALPLPVDPLSKVSTEIPASEPDGPGSVGRAKQDAPLEPT 302
DB 1197 IQGPKESKETPDISTPDSDVAEPLHETIVSE-PAEIQSEEE---IFAQGEYDKLLER 1251
QY 303 F-----HVEITPNVQK-----BOAHSEELHGRAAPFG 329
DB 1252 SDTLQITDLGVSGAREFEVTEPCSEHKGVIESVWTIEDDFITVQVTTTDEGESGHSVRF 1311
QY 330 APEGEPGARGPSLGEDTKADLPESPEKQ-----FAAPRGKPVSRVQLKARVMS 380
DB 1312 AALEQPEVERRPSPHDEEFEEFEVEEAAEAQAPKDGSPAPASPEREVA-LSEYKTETVD 1370
QY 381 KSKDGTGSDD-----KKAATSTRSS-AKTLKNRP 408
DB 1371 DYKDETTIDDSIMDADSLWVDOTDDRSIMTEQLETIPKEKAERAEARSSLEKHKRKEP 1430
QY 409 C-----LSPKLPPTGSSDPL-----IQPSSPA--VCPEPP 436
DB 1431 FKTRGRGRISTPERKVAKEPSTVSRDEVRRKKAVYKAEALAKKTEVQAHSRPSRKETLKPA 1490
QY 437 ---SSPKHVSSTVSRTCSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478

```

```

DB 1491 IKYTRPTHLSCVRRKTTAAGESALAPSVFKQAKDKVSDGVTKSPKRRSLPRPSSILPP 1550
QY 479 QKG-----QANATRIPAKTP-----PAKTPPSS-----GEPPKSG 509
DB 1551 RRGVSGDRDENSESLNSSISSARRTRSEPIRRACKSGTSTTPTGCTAITGCTPPSYS 1610
QY 510 DRSGYSSPGSGTPTGSRSTRPSLTPPP-----TREPKKVAVVRTPPKSPSSAKSRLOTAP 564
DB 1611 SR----TPGTPGTP-SYPRTPHTPGTPKSAIILVPSEKKVAIIRTPPKSPATPK-QLRLIN 1664
QY 565 VPMPLKNNVSKIGSTENLKHQPGGGKVQIINKKLDLSNVQSCGSKDNKIKHVPGGSVQ 624
DB 1665 QPLPDLKNNVSKIGSTENIKYQPKGGQVIVTKKIDLSH----- 1703
QY 625 IVYKPVDLKSVTKSGSIGLNIHHKPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGG 684
DB 1704 -----VTSKCGSLKNIRHRPGGGRVKIESVKLDFEKVQAKVGSLDNAHHVPGGG 1753
QY 685 NKIETHKLTFRENAKARTDHGAIEVYKSPVPGDTSRPHLSNVYSTGSDMVDSPOLAT 744
DB 1754 NVKIDSQKLNFRHAKARVDHGAIEITQPSRSSVASPRRLSNVSSSGSINLLESPOLAT 1813
QY 745 LADEVASASLAKOGL 758
DB 1814 LAEDVTAALAKOGL 1827

RESULT 10
MAP2_MOUSE
ID MAP2_MOUSE STANDARD; PRT; 1828 AA.
AC P20357;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 2 (MAP 2).
GN MAP2 OR TMAP2.
OS Mus musculus (Mouse).
OC Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89083571; PubMed=3205744;
RA Wang D., Lewis S.A., Cowan N.J.;
RT "Complete sequence of a cDNA encoding mouse MAP2.";
RL Nucleic Acids Res. 16:11369-11370(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89043973; PubMed=3142041;
RA Lewis S.A., Wang D., Cowan N.J.;
RT "Microtubule-associated protein MAP2 shares a microtubule binding
motif with tau protein.";
RL Science 242:936-939(1988).
CC -!- FUNCTION: THE EXACT FUNCTION OF MAP2 IS UNKNOWN BUT MAPS MAY
CC STABILIZE THE MICROTUBULES AGAINST DEPOLYMERIZATION. THEY ALSO
CC SEEM TO HAVE A STIFFENING EFFECT ON MICROTUBULES.
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M21041; AAA39490.1; -
DR PIR; S06467; S06467.
DR PIR; A40115; A40115.
DR MGI; 97175; Mtap2.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.

```


[illegible]


```
Db 317 LPTPEDESSTKDVAPPMEBEIEIVGNDDTTSKKTET-----TLPIKMDLAPPEDVLLTKETE 372
Qy 97 IPEGTTABEAGIGDTPSLEDEAAGHVTPESGKVVQEGFLREPQPGGLSHQNSCMP-- 154
Db 373 L-----APAKGMVISEIEEALAKNDVRSAEI-PVAQETVTVSETEVIAVEVLPSPDIT 426
Qy 155 -----GAPLLPEGPREATROPSPGPEDEGGRHAPELLKHQLGLDLHGQGPPLKAG- 207
Db 427 TLTKDVTLPLEAERPLVTDMTFS-LETEMTLGKETAPTEIN--LG-WAKDMSPLPESEV 482
Qy 208 --GKE-----RPGSKEEVEDRDVEDSPQSPSK----- 236
Db 483 TLGKDVILPETKVAEFNNVTPLSEEVTSVKDMSPSAETAPLAKNADLHSGTELIVDN 542
Qy 237 --ASPAQD-----GR-----PQTAAREAT 254
Db 543 SNAPASDLALPLETKVAIVPIKDKTVQTEEPREDSQLASHQHGQSTVPCTASPE-- 600
Qy 255 SIIPGPAEAGIPLPVDFLSKYST-----EIPASEPDGPGSVGRAGQDAPLETFTHVEITPN 310
Db 601 --PVKAAEQMSTLPIDAPSPLENLEQKETPGQSPSEPCSGVSRQEEAKA--VGVGTGN 654
Qy 311 VQKEQAHSEHLGRAAFPCAGPEGPEARG-----PSLGEDTKKADLPPESEKQPA--- 360
Db 655 DITTPPNKE-----PPSPKAKKPLATTQPAKTSTSKAKTQTSPLKPQAPPT 703
Qy 361 -----AAPRGKP-----VSRVPOLKARMYSKSKDGTGSD--DKKAKTS 396
Db 704 SGLNKKPMSLASGVPAPHRPAATATARPSTLPARDV-KRPITKAEKVAERTSPS 762
Qy 397 TRSSARTLKNRCLSPKLPCTPGSSDPLIQ--PS--SPA-VCPEPPSS-----PKHVSSV 445
Db 763 KPSSAPALKPGKPTTPTVSKATSPSTLVSTGSPSSRPATTLPKRPTSIKTEGKPADVKRM 822
Qy 446 TSRTSGGAKEMKLGADGKTKIATPRGAAPPGQKGQANATRIPAKTPPAKTPPSSGEP 505
Db 823 TAKSADLSRSKITTASSVKKRNTTPTGAAPP--AGMTSTRV-----KPMASAPRSSG 873
Qy 506 PKSGDRSGVSSPGSPGTPGSRSTRPSLTPPTPREPKVAVVTPPKPSAKSLQATAPV 565
Db 874 ALSVDKK-----PT-----STRPSSAPRVSLATT-V 900
Qy 566 PMPDLKNVSKITGSTEENLKHOPGG----- 589
Db 901 SAPDLKSVRSKVGSTENIKHOPGGRAKVEKTEAATTAGKEPNVTKAAGSIASAKP 960
Qy 590 --GKVIINKLIDLNVSGSKGSKDNKIHVPGGSGVQIVYVPVLSKVTSCGSLGNTHH 647
Db 961 PAGKVOIVSKKVSYSHIOSKCVSKDNKIHVPGCGNVQIQNKVDISKVSKCGSKANIKH 1020
Qy 648 KPGGQVEVKSEKLFKDRVQSKIGSLDNITHVPGGGNKKIE 689
Db 1021 KPGGGDVKIESQKLFKEKAQKAVGSLDNVGHFPAGGAVKTE 1062

RESULT 13
MAP4_BOVIN
ID MAP4_BOVIN STANDARD; PRT; 1072 AA.
AC P36225;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Microtubule-associated protein 4 (MAP 4) (Microtubule-associated
protein-U) (MAP-U).
GN MAP4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90338002; PubMed=2380192;
```

```
RA Aizawa H., Emori Y., Murofushi H., Kawasaki H., Sakai H.,
RA Suzuki K.;
RT "Molecular cloning of a ubiquitously distributed microtubule-
associated protein with Mr 190,000.";
RL J. Biol. Chem. 265:13849-13855(1990).
RN [2]
RX DOMAINS.
RX MEDLINE=91236765; PubMed=2033072;
RA Aizawa H., Emori Y., Mori A., Murofushi H., Sakai H., Suzuki K.;
RT "Functional analyses of the domain structure of microtubule-associated
protein-4 (MAP-U).";
RL J. Biol. Chem. 266:9841-9846(1991).
CC -1- FUNCTION: NON-NEURONAL MICROTUBULE-ASSOCIATED PROTEIN. PROMOTES
MICROTUBULE ASSEMBLY.
CC -1- TISSUE SPECIFICITY: IS DISTRIBUTED UBQUITOUSLY AMONG ALL TISSUES
BUT AMOUNTS ARE LOWER IN CEREBELLUM AND LIVER.
CC -1- PTM: PHOSPHORYLATION OF THE PRO-RICH REGION IN THE C-TERMINUS
NEGATIVELY REGULATES MAP-4 ACTIVITY TO PROMOTE MICROTUBULE
ASSEMBLY.
CC -1- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90149; BAA14179.1; -
DR PIR; A37127; A37127.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 3.
KW Microtubules; Repeat; Phosphorylation.
FT DOMAIN 244 530 19 X 14 AA TANDEM REPEATS.
FT REPEAT 244 257 1.
FT REPEAT 258 271 2.
FT REPEAT 272 285 3.
FT REPEAT 286 299 4.
FT REPEAT 300 313 5.
FT REPEAT 314 327 6.
FT REPEAT 328 341 7.
FT REPEAT 342 355 8.
FT REPEAT 384 391 9 (INCOMPLETE).
FT REPEAT 392 405 10.
FT REPEAT 406 417 11.
FT REPEAT 418 431 12.
FT REPEAT 432 445 13.
FT REPEAT 446 460 14.
FT REPEAT 461 474 15.
FT REPEAT 475 488 16.
FT REPEAT 489 502 17.
FT REPEAT 503 516 18.
FT REPEAT 517 530 19.
FT REPEAT 907 937 TAU/MAP MOTIF 1.
FT REPEAT 938 968 TAU/MAP MOTIF 2.
FT REPEAT 969 1000 TAU/MAP MOTIF 3.
SQ SEQUENCE 1072 AA; 111914 MW; E8C17A730989F0D2 CRC64;

Query Match 13.7%; Score 545; DB 1; Length 1072;
Best Local Similarity 26.5%; Pred. No. 1.8e-14;
Matches 199; Conservative 109; Mismatches 273; Indels 170; Gaps 29;

Qy 38 DTDAGLKE-PLQPTTEDSGSEPGSETSDAKSTP-----TAEVDVTAPLVDEGAPGKQA 89
Db 325 ETDVSLDEDLALSTETETVSTAQDILLKFKETESIPPTMDLSAEGTVPPTDQEMTPKVA 384
Qy 90 AAQPHTEIP-----EGTTAEAGIGTGTSPSLDEAAGHVTPESGKVVQEGFLREP 140
Db 385 ASLSEIEAPLDEEDIVSSTEIRSAKEIGL-----SSETEVA----- 419
Qy 141 GPGGLSHQLMSGMPGAPLLPEGPPEATRPQSGCTGPEDEGGRHAPELLKHQLGLDLHQEC 200
```



```
Db 420 ----LSREM-----GLP-----PE-----TEAIDKDMAAPPETE-----VIVPVKDMA 454
Qy 201 PPLKGAGGKERGSKEEVEDD-----RDVDESSPODS-----PPSKASP 239
Db 455 P-----SPGIETTLAKOVAPRQEEVLTKGDTVSLPATEMALGRNALVPETETVTL 505
Qy 240 AOD-GRPPQTAAREA--TSIPGFPAEAGAIPLVPDLSFKVSTFIPASEPDGPGSVGRAGQD 296
Db 506 AKDVAQPPEVETNLANNAALAKFSEAEVVPVKDMETAQTQEATSEDSQLKSLQDEQS 565
Qy 297 ABLETFHVEITPNVOKQAHSEHGLRAAPGAPGEGPGLSGDTKEADLPDPSE 356
Db 566 AVPLMTSPAVVAMQKHSPLTDEDD---SVLEELBQKKPSSQTSPLPSET---SGVAKPEE 620
Qy 357 KOPAAAPRGKPVSRVLPOLKARVMSKDGTSDDKKKTSRSSAKTLKLRPCLSPKLP 416
Db 621 GPPTGSVSGNDITAPPNKELPPSPPEKTKPLATTPQAKTST-SKAKT---QPTSLPKQTA 676
Qy 417 P---GSSDPLIOPSPPAVCPEPPSPKPVSVTSRTSGSGAKMKLG-ADGK---TKIA 469
Db 677 PITLGSNK--KPSLASSGVPAAPKPPAAATSRPSTLPSKDTKPKPVAEKIPKRV 734
Qy 470 TPRGAAPGQKQGANATRIAPKTP-----PAKTP---PSSGEPPKSG 509
Db 735 PSKPASAPAVKPGSKSTQAVKAPATATLASPGSTSRNLSTPLPKRPTAIKTEGKPAEIK 794
Qy 510 DRSGYSSGSGPCTPGS-----RSRTSLPTPTPREKKVAVVTPPK-----551
Db 795 KWATSPADLRSKSTTTSSVSKSTTVPGTAPPAGASRPATPRPGTGPVDDKPP 854
Qy 552 ---SPSSAKSRL--QTAPVPMPLKNVSKIGSTENLKHQPGGKVQIINKLKLDS-----602
Db 855 TAAKPTSSAPRLGRVAANASAPDLANVRSKVSTENIKHQPGGGRK-VEKTEAAAPAR 913
Qy 603 ---NVQSKGSKDNKHVPGGGSVQIVKPVDSLVKTSKCSGLNIHHKPGGGGVEVKS 658
Db 914 KPENAVTKAAGPIGNAQKPTGKVOIQNKVKYDISKSKGSKRANIKHKPGGGVDKIES 973
Qy 659 EKLDKDRVQSKIGSLDNITHVPGGNGKKIE 689
Db 974 QKLNFEKAQAQVGLSDNVGHLPAGGAVKTE 1004

RESULT 14
TCOF HUMAN
ID TCOF HUMAN STANDARD; PRT; 1411 AA.
AC Q13428; Q99408; Q99860;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Treacle protein (Treacher collins syndrome protein).
GN TCOF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154183; PubMed=8563749;
RA Dixon J., Edwards S.J., Gladwin A.J., Dixon M.J., Loftus S.K.,
RA Bonner C.A., Koprivnikar K., Wasmuth J.J.;
RT "Positional cloning of a gene involved in the pathogenesis of
RT Treacher Collins syndrome."
RL Nat. Genet. 12:130-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97228900; PubMed=9074926;
RA Dixon J., Edwards S.J., Anderson I., Brass A., Scambler P.J.,
RA Dixon M.J.;
RT "Identification of the complete coding sequence and genomic
RT organization of the Treacher Collins syndrome gene."
RL Genome Res. 7:223-234(1997).
```

[3]
SEQUENCE FROM N.A.
MEDLINE=97250498; PubMed=9096354;
Wise C.A., Chiang L.C., Paznekas W.A., Sharma M., Musy M.M.,
Ashley J.A., Lovett M., Jabs E.W.;
"TCOF1 gene encodes a putative nucleolar phosphoprotein that exhibits
mutations in Treacher Collins syndrome throughout its coding
region.";
Proc. Natl. Acad. Sci. U.S.A. 94:3110-3115(1997).
[4]
VARIANTS L-439; V-810; V-1313 AND G-1355, AND VARIANT TCS R-53.
MEDLINE=97195537; PubMed=9042910;
Edwards S.J., Gladwin A.J., Dixon M.J.;
"The mutational spectrum in Treacher Collins syndrome reveals a
predominance of mutations that create a premature-termination
codon.";
Am. J. Hum. Genet. 60:515-524(1997).
-1- DISEASE: DEFECTS IN TCOF1 ARE THE CAUSE OF TREACHER COLLINS
SYNDROME (TCS). TCS IS A AUTOSOMAL DOMINANT DISORDER OF
CRANIOFACIAL DEVELOPMENT THAT OCCURS WITH AN INCIDENCE OF 1/50,000
LIVE BIRTHS. THE CLINICAL FEATURES OF TCS ARE BILATERALLY
SYMMETRICAL AND INCLUDE: (1) ABNORMALITIES OF THE EXTERNAL EARS,
ATRESIA OF THE EXTERNAL EAR CANALS, AND MALFORMATION OF THE MIDDLE
EAR OSSICLES, WHICH MAY RESULT IN CONDUCTIVE HEARING LOSS; (2)
LATERAL DOWNWARD SLOPING OF PALPEBRAL FISSURES, FREQUENTLY WITH
COLOBOMAS OF THE LOWER EYELIDS; (3) HYPOPLASIA OF THE MANDIBLE AND
ZYGOMATIC COMPLEX; (4) CLEFT PALATE.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U40847; AAC50903.1; --
EMBL; U76366; AAC51181.1; --
EMBL; U84664; AAC51185.1; --
EMBL; U84640; AAC51185.1; JOINED.
EMBL; U84641; AAC51185.1; JOINED.
EMBL; U84642; AAC51185.1; JOINED.
EMBL; U84643; AAC51185.1; JOINED.
EMBL; U84644; AAC51185.1; JOINED.
EMBL; U84645; AAC51185.1; JOINED.
EMBL; U84646; AAC51185.1; JOINED.
EMBL; U84647; AAC51185.1; JOINED.
EMBL; U84648; AAC51185.1; JOINED.
EMBL; U84649; AAC51185.1; JOINED.
EMBL; U84650; AAC51185.1; JOINED.
EMBL; U84651; AAC51185.1; JOINED.
EMBL; U84652; AAC51185.1; JOINED.
EMBL; U84653; AAC51185.1; JOINED.
EMBL; U84654; AAC51185.1; JOINED.
EMBL; U84655; AAC51185.1; JOINED.
EMBL; U84656; AAC51185.1; JOINED.
EMBL; U84657; AAC51185.1; JOINED.
EMBL; U84658; AAC51185.1; JOINED.
EMBL; U84659; AAC51185.1; JOINED.
EMBL; U84660; AAC51185.1; JOINED.
EMBL; U84661; AAC51185.1; JOINED.
EMBL; U84662; AAC51185.1; JOINED.
EMBL; U84663; AAC51185.1; JOINED.
EMBL; U79659; AAB40722.1; JOINED.
EMBL; U79645; AAB40722.1; JOINED.
EMBL; U79646; AAB40722.1; JOINED.
EMBL; U79647; AAB40722.1; JOINED.
EMBL; U79648; AAB40722.1; JOINED.
EMBL; U79649; AAB40722.1; JOINED.
EMBL; U79650; AAB40722.1; JOINED.
EMBL; U79651; AAB40722.1; JOINED.
EMBL; U79652; AAB40722.1; JOINED.
EMBL; U79653; AAB40722.1; JOINED.


```

DR EMBL; U79654; AAB40722.1; JOINED.
DR EMBL; U79655; AAB40722.1; JOINED.
DR EMBL; U79656; AAB40722.1; JOINED.
DR EMBL; U79657; AAB40722.1; JOINED.
DR EMBL; U79658; AAB40722.1; JOINED.
DR Genew; HGNC:11654; TCOF1.
DR MIM; 606847; -.
DR MIM; 154500; -.
DR InterPro; IPR003993; treacle.
DR Pfam; PF03546; treacle; 3.
DR PRINTS; PR01503; TREACLE.
KW Disease mutation; Polymorphism.
FT DOMAIN 89 97
FT POLY-GLU.
FT DOMAIN 204 207
FT POLY-SER.
FT DOMAIN 616 619
FT POLY-SER.
FT DOMAIN 919 924
FT POLY-SER.
FT DOMAIN 1285 1289
FT POLY-LYS.
FT DOMAIN 1375 1386
FT POLY-LYS.
FT DOMAIN 1398 1405
FT POLY-LYS.
FT VARIANT 53 53
FT W -> R (IN TCS).
FT /FTID=VAR_005630.
FT VARIANT 439 439
FT P -> L.
FT /FTID=VAR_005631.
FT VARIANT 810 810
FT A -> V.
FT /FTID=VAR_005632.
FT VARIANT 1313 1313
FT A -> V.
FT /FTID=VAR_005633.
FT VARIANT 1355 1355
FT D -> G.
FT /FTID=VAR_005634.
FT CONFLICT 1312 1312
FT K -> Q (IN REF. 2).
FT SEQUENCE 1411 AA; 144312 MW; 3880203D985C2699 CRC64;

Query Match 7.5%; Score 297; DB 1; Length 1411;
Best Local Similarity 21.1%; Pred. No. 6.2e-05;
Matches 164; Conservative 92; Mismatches 279; Indels 242; Gaps 33;

QY 1 MAEPFQEFVEMDHAGTYGLGRKDGQGYTMHQDQEGDTPDAGLK-----ESPLOTPTED 54
DB 40 LAQPTLLDIYTHQQTSELGRKR-----KAEDDALQAKKTRVSDPST-SES 87
QY 55 GSEEPGSESDAKSTTAEVDVAPLVEGAP-----GKQAAQPHFTPEEG 100
DB 88 SEEEAEAEATKATPRLASTNSVLGADLPSSMKREKAKAETKAGKTGNSMPHP--ATG 145
QY 101 TTAEEAGIGDTSLEDAAGH---VTOEPESCKVQVQEGFLREPQPPG----- 144
DB 146 KTVANLLSKSPKSNPSANTLTVSETEEGVSFAFGAAKPGMVASQAGDSSSEDTS 205
QY 145 -----LSHQLM-----SGMPGAPLLPEGPPE---ATROPSTGTGPDTEGG 181
DB 206 SSETDVEVKASEKILQVRAASAPAKGTGKATPAPPGKAGAVASQTKAGKPEEDSES 265
QY 182 RH-----APLLKHQLLDLHQBG-----PPLKAG-----GKERP----- 212
DB 266 SESSDSEETPAAKALLOKASGKTSQVGAASAPAKESPRGGAAPPPGKTGPVAKAK 325
QY 213 GKKEEVEDRDVDESSPQSPKSPASQADGRPPOTAAREATSIQFPAEGAIPLPVDFL 272
DB 326 AKGREDSOSSSEESDEEAEPAQAKPS--GKAPQVRAASAPAKES-PRKGAAPPAPPKT 382
QY 273 SKVSTEIPASEPDGPGSVGRAGQDAPLEFTFIVEI-----TPNVQEQAHSEHLGR 324
DB 383 GPAAQVQGVQKQEDSSRSSEESDSDREALAAMNAQVKPLGKSPQVRPASTMGWGLPK 442
QY 325 AAFPAGPGE-GPEARGPSLGE-----DTKEADLP---EPEK-----Q 358
DB 443 GAGVPVPGVGPATPSAQVQKWEEDSESSSESSSDSGEVPVTAVAPAEKSLGNLOAK 502
QY 359 PAAAP--RGRPVSRVP---OLKA-----RMVSKSKDGTGSDDKA----- 393
DB 503 PTSSPAKGPQKAGPVAVQVQKAEKPMWNSSESSSDSADSEAEPAAMTAQAKPALKIP 562
QY 394 -----KTSTRSSAKTLKNRCLSPKLPTPGSDPLIQSPSPAVC-----PEPPSPK 440

```

```

Db 563 QTKACPKKTKNTASAKVAPVAVGTPPRKAGTATSP--AGSSPAVAGGTQTPADSSSE 620
QY 441 HVSSVTSRTG-----SSGAKEMKLGADGKTATPRGAAPPQKQGANATRIPAKT 492
Db 621 ESDSEEXTGLAVTVGQAKSVGKGLQVKAASVPVK-----GSLGCGTAPVLPCKT 670
QY 493 PP-----APKTPPSSGEPKSGDRSGYSSP-----GSPG 521
Db 671 GPTVTQVKAKEODESESESESESESESESESESESESESESESESESESESESESE 730
QY 522 TPG-----SRSTPSLPTPTREPKKVAV--RTPPKSPSSAK-----SRLQTP 564
Db 731 APGVVTTAAQAQKQSPKSVKVPVNRPNPONTSTVLARGPASVPSVGVKAVATAQAOTGP 787

RESULT 15
CAL3_MOUSE
ID CAL3_MOUSE STANDARD; PRT; 1464 AA.
AC P08121; Q61429; Q9CRN7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X DBA; TISSUE=Embryo;
RX MEDLINE=95011609; PubMed=7926795;
RA Tonan D., de Crombrughe B.;
RT "The mouse type-III procollagen-encoding gene: genomic cloning and
RT complete DNA sequence.";
RL Gene 147:161-168(1994).
RN [2]
RP SEQUENCE OF 1-488 FROM N.A.
RX MEDLINE=88167858; PubMed=3443309;
RA Wood L., Theriault N., Vogeli G.;
RT "Complete nucleotide sequence of the N-terminal domains of the murine
RT alpha-1 type-III collagen chain.";
RL Gene 61:225-230(1987).
RN [3]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=85131189; PubMed=3972847;
RA Liau G., Mudryj M., de Crombrughe B.;
RT "Identification of the promoter and first exon of the mouse alpha 1
RT (III) collagen gene.";
RL J. Biol. Chem. 260:3773-3777(1985).
RN [4]
RP SEQUENCE OF 810-1464 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;

```


"Functional annotation of a full-length mouse cDNA collection.";

```

RT Nature 409:685-690(2001).
RN [5]
RC SEQUENCE OF 1442-1464 FROM N.A.
RX MEDLINE=91274355; PubMed=2054384;
RA Metsaeranta M., Toman D., de Crombrughe B., Vuorio E.;
RT "Specific hybridization probes for mouse type I, II, III and IX
RL collagen mRNAs";
RL Biochim. Biophys. Acta 1089:241-243(1991).
CC -!- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
CC ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC O-LINKED GLYCAN CONSISTS OF GLC-GAL DISACCHARIDE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52046; CAA36279.1; -
DR EMBL; M18933; AAM37338.1; -
DR EMBL; K03037; -; NOT_ANNOTATED_CDS.
DR EMBL; AK019448; BAB31724.1; -
DR EMBL; X57983; CAA41048.1; -
DR PIR; A22287; A22287.
DR PIR; A27353; A27353.
DR PIR; S16373; S16373.
DR MGD; MGI:88453; Col3a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib.collagen_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF01391; Collagen; 18.
DR Pfam; PF01410; COLFI; 1.
DR ProDom; PD000007; Collagen; 1.
DR ProDom; PD002078; Fib.collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01208; VWFC; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT PROPEP 24 154 AMINO-TERMINAL PROPEPTIDE.
FT CHAIN 155 1203 COLLAGEN ALPHA 1(III) CHAIN.
FT PROPEP 1204 1464 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN 31 90 VWFC.
FT DOMAIN 155 169 NONHELICAL REGION (N-TERMINAL).
FT DOMAIN 170 1195 TRIPLE-HELICAL REGION.
FT DOMAIN 1196 1464 NONHELICAL REGION (C-TERMINAL).
FT CARBOHYD 262 262 O-LINKED (GAL...) (BY SIMILARITY).
FT MOD_RES 262 262 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 283 283 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 859 859 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 976 976 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1093 1093 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1105 1105 HYDROXYLATION (BY SIMILARITY).
FT DISULFID 1195 1195 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1196 1196 INTERCHAIN (BY SIMILARITY).
FT SEQUENCE 1464 AA; 138944 MW; 2104EC27A886090B CRC64;
Query Match 7.4%; Score 294; DB 1; Length 1464;
Best Local Similarity 23.2%; Pred. No. 8.3e-05;
Matches 165; Conservative 52; Mismatches 249; Indels 244; Gaps 33;
QY 7 EFWMEDHAGTYGLGRDKDGGYTHMQDQGGYTHMQDQGGTDAGLKESPLQTPTEDEGSE-EPGSETSD 65

```

```

Db 155 QFSDYDVKSGVGMGYPPGAG-----PPGPPGSSGHPGS----- 192
QY 66 AKSTPTAEDVTAPLVDEGAPCKQA-AAQPHTEIPEGTTAEAGIGDTPTSLDEDAAGHVQTQ 124
Db 193 -----PGSPGYQGPGEPPGAGPPGPPGALG-----PAGPAGK 228
QY 125 EPESGKVVQEGFLREPQPPGLSHOLMSGMPGAPLL----- 159
Db 229 DGESEGRPCRERGLPGPPGLKGP-AGMPCFFPMKCHRGFDGRNGKEGTGAPLKG 286
QY 160 -----PEGPREATRQ-----PSGTGPEDETEGRHAPELLKHQLGLDHLHGEPP 203
Db 287 GLPGDNGAPGPMGRGAPGERGRGLPCAAGARGNDGAR-----GSDGQCPGP- 334
QY 204 KG-----AGGRERPCSKKEVEDRDVDESSPDSPSKASPAQDGRPPQTAAREATSTPGF 259
Db 335 -GPPTAGFPGSGAKGEV-----GAGSPGSGPQGRGEPGQGHAGAGQPPGP 384
QY 260 PAEGAIPLPVDFLSKVSTEIPASEPDGSPVGRARQDAPLETFTHVEITPNVQKEQAHSE 319
Db 385 PGNNSGPGKGEVG-----PAGIPGAPGLIGARGPPGAG-----TNGIPCTRGPSG 431
QY 320 E--HLGRAAPGAPGEGPEARGPSL-----CEDTKHEADLPEP-----SEKQPA----- 360
Db 432 EPGKNGAKGEFGARGERGEAGSGIPGPKGEDKDGSPGEPGANGLPAAAGERGPGSGFRG 491
QY 361 -AAPRGKPVSRVPOLK-----ARMVSKSDGTGSDDKKA 393
Db 492 PAGNNGIPGEKGPGERGGPGPAGPRGAVAGEPRDGTGPGGIRGMGPGSGGGNDGKPG 551
QY 394 KTSTRSSAKTLKNNRPLCLSPKLTPTGSSDPLIQPSSPAVCPPEPPSPKHVSVTSRTGSSG 453
Db 552 PPGSGGES-----GRP-----GPPGSGPRGQ-----PCVMGFP--GPKGNDGAPCKNGERG 596
QY 454 AK-EMKLGADGKTKIATPRG-----AAPQCKQANATRIAPKTPPAK-----T 498
Db 597 GPGGFGLPFGPAGKNGETGPGQPPGPTGPDGKDSG-----PPGPGLOQIPCTGPG 648
QY 499 PPSGGEPPKSGDRSGSYSGSPGSP-----GTPGSR--RTPSLP-----TPPTREPCKVA 544
Db 649 PGENGKFGEPGPKGEVAGAPGAPGCKGDSGAPGERGPGTAGIPGARGGAGPPGPGGKGP 708
QY 545 VVTRTPPKSPSSAKSRLQTAPEVPMPLDNKVKSKIG-----STENLKHQPGG 589
Db 709 A--GPPGPPGASGS-----PGLQGMPCERGSGSPGPKGKEGEPGG 747

```

Search completed: January 28, 2003, 14:15:17

Job time : 26.5946 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 38.3037 Seconds

(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPRQEFVMEHDHAGTYGL.....SPQLATLADEVSAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp Vertebrate.*
14: sp Unclassified.*
15: sp_virus.*
16: sp_bacterioph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1537	38.8	372 11 Q91WK4	Q91WK4 mus musculus
2	795	20.1	1825 11 Q64715	Q64715 rattus norv
3	790.5	19.9	1828 11 Q63724	Q63724 rattus ratt
4	715	18.0	323 6 Q28869	Q28869 bos taurus
5	704	17.8	801 6 Q95KD1	Q95KD1 macaca fasc
6	690.5	17.4	198 11 P97749	P97749 rattus sp.
7	576	14.5	381 11 Q64710	Q64710 rattus norv
8	567.5	14.3	414 11 Q9DBX9	Q9DBX9 mus musculus
9	485	12.2	838 5 Q8T5L5	Q8T5L5 anopheles g
10	482.5	12.2	1224 13 Q9YI90	Q9YI90 xenopus lae
11	481	12.1	125 4 Q16296	Q16296 homo sapien
12	471	11.9	700 5 Q9VB15	Q9VB15 drosophila
13	392	9.9	102 11 Q9QVU4	Q9QVU4 rattus sp.
14	379.5	9.6	361 11 Q95UT1	Q95UT1 drosophila
15	377.5	9.5	453 5 O02592	O02592 caenorhabdi
16	371	9.4	928 13 Q98906	Q98906 gallus gall

17	362	9.1	431	5	Q17364	Q17364 caenorhabdi
18	355	9.0	102	11	Q9QVU3	Q9QVU3 rattus sp.
19	350	8.8	339	4	Q8TB56	Q8TB56 homo sapien
20	347	8.8	113	13	Q12967	Q12967 fugu rubrip
21	342.5	8.6	413	5	P90973	P90973 caenorhabdi
22	341	8.6	458	5	Q95OG4	Q95OG4 caenorhabdi
23	325.5	8.2	436	5	Q17365	Q17365 caenorhabdi
24	309	7.8	2187	11	P70670	P70670 mus musculu
25	299.5	7.6	5120	13	Q9PU36	Q9PU36 gallus gall
26	296	7.5	958	5	Q8SXI5	Q8SXI5 drosophila
27	294.5	7.4	1422	6	Q95KU4	Q95KU4 canis fami
28	294	7.4	958	4	Q96A52	Q96A52 homo sapien
29	289	7.3	885	5	Q9TXR9	Q9TXR9 caenorhabdi
30	289	7.3	930	5	Q17339	Q17339 caenorhabdi
31	288.5	7.3	847	6	Q95KP7	Q95KP7 canis fami
32	285.5	7.2	990	13	Q91803	Q91803 xenopus lae
33	285	7.2	971	5	Q9XVS4	Q9XVS4 caenorhabdi
34	285	7.2	1182	4	Q99495	Q99495 homo sapien
35	285	7.2	1835	13	Q91AU4	Q91AU4 gallus gall
36	284.5	7.2	1820	13	Q91907	Q91907 pagrus majo
37	284.5	7.2	2944	11	Q63870	Q63870 mus musculu
38	283.5	7.2	1745	4	Q9NZQ6	Q9NZQ6 homo sapien
39	283	7.1	1838	4	Q15094	Q15094 homo sapien
40	282	7.1	1151	13	O57580	O57580 gallus gall
41	282	7.1	1450	13	Q9YIB4	Q9YIB4 cynops pyrr
42	282	7.1	2703	5	Q9VEG7	Q9VEG7 drosophila
43	281.5	7.1	1445	13	Q93251	Q93251 rana catesb
44	281	7.1	1458	13	Q910B9	Q910B9 oncorhynch
45	281	7.1	1838	11	O88207	O88207 mus musculu

ALIGNMENTS

RESULT 1

Q91WK4
ID Q91WK4 PRELIMINARY; PRT; 372 AA.
AC Q91WK4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein tau.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE, AND RETINA;
RA Strausberg R.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014748; AAH14748.1; -
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 4.
DR PROSITE; PS00229; TAU_MAP; UNKNOWN 4.
SQ SEQUENCE 372 AA; 3861 MW; B027745D23BC62A2 CRC64;

Query Match 38.8%; Score 1537; DB 11; Length 372;
Best Local Similarity 45.3%; Pred. No. 3.9e-79;
Matches 344; Conservative 12; Mismatches 14; Indels 390; Gaps 6;
Qy 1 MAEPRQEFVMEHDHAGTYGLGDRKDDQGYTHQDOEGDAGLKESPLQPTTDEGSEEPG 60
Db 1 MADPRQEFDTMEDHA-----GDYTLQDQEGDMDHGLK----- 33
Qy 61 SETSDAKSTPTAEDVTPALVDGAPGKQAAQAHPTEPGTAAEAGTGDTPSLEBAAG 120
Db 34 -----AEEAGIGDTPNDEQAAG 51
Qy 121 HVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 52 HVTQ----- 55


```
QY 181 GRHAPELLKHQLLDHQBGPPLKGAGGKERPGSGKEEVEDDRDVEDSSPDQSPSPKASPA 240
Db 56 -----
QY 241 QDGRPPQTAAREATSIPIGPAEGAIPLPVDFLSKVSSTEIPASEPDGSPVGRAGQDAPLE 300
Db 56 -----ARVA-----
QY 301 FTFHVITPNVQEAHSEHLGRAFFAGPAGGEPGARGPSLGEDTKEADLPPESEKQPA 360
Db 60 -----
QY 421 DPLIQSSPAVCPPEPSSPKHYSSVTSRTGSSGAKEMKLKGADGKT--KIATPRCAAPP 478
Db 72 -----KGADGKTGAKIATPRGAASPA 92
QY 479 QKQANATRIAPKTPAPKTPPSSGEPKPSGDRSGYSPPGTPGCSRSRTPSLTPPTR 538
Db 93 QKGTSNATRIAPKTPSPKTPGSGEPKPSGDRSGYSPPGTPGCSRSRTPSLTPPTR 152
QY 539 EPKKVAVRTPKPSASRLQATAPVMPDLKNVSKIGSTENLKHOPGGGKVOIINKK 598
Db 153 EPKKVAVRTPKPSASRLQATAPVMPDLKNVSKIGSTENLKHOPGGGKVOIINKK 212
QY 599 LDLSNVQSKGSKDNKHKHVPGGGSGVOIVKPYDLSKVTSCGSLGNIHHKPGGGQVEVKS 658
Db 213 LDLSNVQSKGSKDNKHKHVPGGGSGVOIVKPYDLSKVTSCGSLGNIHHKPGGGQVEVKS 272
QY 659 EKLDKDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
Db 273 EKLDKDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 332
QY 719 DTSRHLNSVSTGSDIMVDSQPLATLADSVASLAKQGL 758
Db 333 DTSRHLNSVSTGSDIMVDSQPLATLADSVASLAKQGL 372

RESULT 2
Q64715 PRELIMINARY; PRT; 1825 AA.
AC Q64715;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Microtubule-associated protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RA Dresse A.E.;
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1802-1825 FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RX MEDLINE=89334524; PubMed=2474284;
RA Marechal D., Delapierre D., Dresse A.;
RT "Cloning and partial sequencing of a new rat brain specific cDNA.";
RL Arch. Int. Physiol. Biochim. 96:231-236 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RX MEDLINE=95094034; PubMed=8000911;
RA Ferhat L., Ben-Ari Y.P., Khrestchatsky M.;
RT "Complete sequence of rat MAP2a, a novel MAP2 isoform.";
RL C. R. Acad. Sci., III, Sci. Vie 317:304-309 (1994).
CC -1- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
```

```
DR EMBL; X54100; CAA38034.1; -.
DR EMBL; X74211; CAA52283.1; -.
DR IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.
KW Microtubules; Repeat.
FT CONFLICT 151 1509 MISSING (IN REF. 3).
FT CONFLICT 1689 1689 Q -> VRILNKKIDFSKVSRCGSKDNKIKHAGGN (IN
FT REF. 3).
SQ SEQUENCE 1825 AA; 198565 MW; 6A920FD9B22213BC CRC64;

Query Match 20.1%; Score 795; DB 11; Length 1825;
Best Local Similarity 29.2%; Pred. No. 1.6e-36;
Matches 275; Conservative 118; Mismatches 273; Indels 276; Gaps 34;

QY 3 EPROEFVEFMEHDAGTYGLG-----DRKDOGGYTMHODQEGDT---- 39
Db 974 EHAKSEVGDGVKVELFGLGVYEQTSKALITTKETAPERAEGKLSVPEVAEVTTKA 1033
QY 40 DAGL-----KESPLQTPTEGSEEPGSETDAKST-----PTAED- 74
Db 1034 DOGLDVAAKDDQSPLDIKVSDFGOMASGMSVDAGKTIELKFEVDQQLTSLSEAPQETDS 1093
QY 75 ---VTAPLVDEGA-----PGQAAAQPHTEIPGTTABEA--GIGOTPSLEDA----AG 120
Db 1094 FMGIESHVKGAKVSETEVKVAKPDLVHQEAVDKEESYESSGESHESLTMSLRKPEDG 1153
QY 121 HVTQEPESGKVQVQEGFLR---EPGPPGLSHQLMS-----GMP-----GAPLL 159
Db 1154 KKETSPETSLIODEVALKLSVEIPCPPVSEADSIDAEKAEVQMEFIQLPKESTETPDI 1213
QY 160 PEGPREATR-QPSGTGPDTEGGRHAPPELLKHLQLLDHQBGPPLKGAGKER----- 211
Db 1214 PAIPSDVTQPOEAVVSEPAE-----VRGEEIEAEGEVDKLLPRSDT 1257
QY 212 -----PGSKEEV-----DEDRDVDES-----SPQDSPSKASPAQGRPPQTAAREATS 255
Db 1258 LQITDLVPGSREEFVETCPGEHGVVESVVTIEDDTITVVTQTTDEGLGSHSVRFAA- 1316
QY 256 IPGPPAEGAIPLPVDFLSKV--STEIPASEPDGSPVGRAGQDAPLETFTHVEITPNVQK 313
Db 1317 -PVQPEERRYPHDEEVLVMAEAQAEPKDGSDAPATPEKEEVFSEYKETYDYDK 1375
QY 314 EQAHSEHLGRA-----APPGAPGEGPEARGPSLGEDTKEADLPE 353
Db 1376 DETTIDDSIMDADSLWVDOTDDRSILTEQLETTIPKEERAKEARRPSLEKHKRKE----- 1430
QY 354 PSEKQPAAPRGKPVSRVPQLKARWVSKSDGTGSD-----KKATSTR 398
Db 1431 ----KPKFTGRGR--ISTPE---RKVAKKEPSTVSRDEVRKKKAVYKKAELAKKSEVOAH 1481
QY 399 SSAKTLKNRPCLSPKLPDTPGSSDPLIQPSSPAVCPPEPSSPKHYSSVTSRT----- 449
Db 1482 SPSKRLILKPAIKYTRPT-----HLSCVKRKTATSGESAQ 1517
QY 450 GSSGAKEMKLKGADGKTKIATPRGAAP-----PGQKG-----QANATRIAPKTPPAPKT 498
Db 1518 APSAFKQAKDKVTDGITSKPEKSSLPSPSSILPRRGSGDRRENSFSLNSSISSARRT 1577
QY 499 PPSGEPKPSGDRSGYSPPGSG-----TPGS-----RSRTPSLTPPTT--REP----- 540
Db 1578 TRS---EPIRRAGKSGTSTPTTPGSTAITPCTPPPSYSRTPTGCTGTPSYRPTGPKFGIL 1635
QY 541 ----KVAVVRTPKPSASSAKSLQATAPVMPDLKNVSKIGSTENLKHOPGGGKVOIIN 596
Db 1636 VPSEKKVAIITPPKSPATPK-QRLINQPLDPLKNVSKIGSTDNIKYQPKGGQVOIVT 1694
QY 597 KKLDSNVQSKGSKDNKIKHVPGGGSGVOIVKPYDLSKVTSCGSLGNIHHKPGGGQVEV 656
Db 1695 KKIDLSH-----VTSCGSLKNIRHPPGGGRVKI 1723
QY 657 KSEKLDKDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVKSPV 716
Db 657 KSEKLDKDRVQSKIGSLDNITHVPGGSKKIETHKLTFRENAKAKTDHGAIEIVKSPV 716
```



```
Db 1724 ESVKLDFEKAQAKVGSGLDHAHVPGGNVKIDSKLNFREHAKARVDHGAIBITQSPSR 1783
QY 717 SGTSPRLHNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
Db 1784 SSVASPRRLNSVSSGSINLLESPQLATLAEDVTAALAKOGL 1825

RESULT 3
Q63724
ID Q63724 PRELIMINARY; PRT; 1828 AA.
AC Q63724;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10117;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=RA25; TISSUE=BRAIN;
RA Matus A., Doll T.;
RL Submitted (MAY-1990) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
DR EMBL; X53455; CAA37535.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.
KW Microtubules; Repeat.
SQ SEQUENCE 1828 AA; 198983 MW; 553C3BB5F16D436C CRC64;

Query Match 19.9%; Score 790.5; DB 11; Length 1828;
Best Local Similarity 29.7%; Pred. No. 2.8e-36;
Matches 276; Conservative 121; Mismatches 280; Indels 253; Gaps 35;

QY 3 EPRQEFVEMDHAGTYGLG-----DRKQGGYTMHQDEGDT---- 39
Db 978 EHAKSEEVGDKVFLGLGVTYEQTSAKELITTKETAPRAEKGGLSSVPEVAEVETTKA 1037
QY 40 DAGL-----KESPLQTTPEGSEPGSETSDAKST-----PTAED- 74
Db 1038 DQGLDVAARKDQSDPLDKVDFGQWASGMSVDAGTKIELKFEVDQQLTSLSEAPQETDS 1097
QY 75 ---VTAPLYDEGA----PGKAAQOPHTEIPEGITAEAA--GIGDTPSLEDEA---AG 120
Db 1098 FWGIESHVKDKAKVSETVEKVKAPDLVHQEAVDKVEESSEGESHTNESLKPDGEG 1157
QY 121 HVTQEPESGVVQEGFLR----EPGPPGLSHQMSGMPGAPL-----LPEGPREATROP 170
Db 1158 KKETSPTSLSIQDEVALKLSVEIPCPPVPVSEADSSIDEKAEVQMEFIQLPK--EESTETP 1215
QY 171 SGTG-PEDTEGGHAPPELLKHLLGDLHQEGPLPKAGGKER-----PGS 214
Db 1216 DIPAIPSDVT--QPOEALVSE--PAEVRGEEEEEIEAGEYDKLLFRSDTLQITDILLVPGS 1272
QY 215 KEEV-----DEDRVDSES---SPQSPSPKSPADQGRPPQPAAREATSIQCFPAEGAIP 266
Db 1273 REEFVETCGEHKGVVSVVTTEDFITVVQTTTDEGELGSHSVRFAA--PVQPEEERRP 1330
QY 267 LPVDFLSKV--STEIPASPDGSPVGRKQGDAPLEFTFHVITTPNVQKEQAHSEHLGR 324
Db 1331 YPHDELEVLMAEAQAEPKDGSPDAPATPEKEEVPFSEYKETTYYDDYKDETTIDDSIMD 1390
QY 325 A-----AFPGAGGEGPEARGPSLGEDTKADLPPEPSKQPAAPR 364
Db 1391 ADSLWVDTODDRSILTEQLETIPKEERAKEARRPSLEKHKR-----KPKFTR 1441
QY 365 GKPVSNVPLKARMYSKSDGTGSD-----KKAATSTESSAKTLKNRCP 409
Db 1442 GR--ISTPE--RKVAKKEPSTVSRDEVRKRVKAVYKKAELAKKSEVQAHSPSKLILKPA 1496
QY 410 LSPKPLTPGSDPLQIPSSPAVCPPEPPSPKHKVSSVTSRT-----GSSGAKEMKILK 460
```

```
Db 1497 IKYTRPT-----HLSVCRKRTTATSGESAQAASAKQAKDK 1532
QY 461 GADGKTKATPRGAAP-----PQKG-----QANATRIPAKTPAPKTPSPSSPEPKSG 509
Db 1533 VTDGITSPEKRSSILPRPSSILPRRGSVDRENSFSLSNLSISSARTRTS--EPIRRA 1590
QY 510 DRSGYSSPGSPG---TPGS---RSRTPSLTPPT--REP-----KKVAVVRT 548
Db 1591 GKSGTSTPTTPTGTAITPTGTPSYSSRTPGTPGTPSPYRPTPTGTPKSGILVPSKVAIIRT 1650
QY 549 PPKSPSSAKSLRLQAPVPMPLKNVKSIGSTENLKHOPGGKQVQIINKKLDLNSVQSKC 608
Db 1651 PPKSPATPK-QLRLINQPLPLKNVKSIGSTDNIKYQPKGQVQIVTKKIDLSH----- 1704
QY 609 GSKDNIKHVPGGGSVQIVKPVDLISKVTSKCSGLNTHHKGPGGQVEYKSKLDFKDRVQ 668
Db 1705 -----VTSKCSGLKNIRHRPGRGGRVKIESVKLDFEKEAQ 1738
QY 669 SKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVVGDTSPRHLSNV 728
Db 1739 AKVGSLDNAHVPGGNGVKIDSKLNFREHAKARVDHGAIEITQSPSRSSVASPRRLSNV 1798
QY 729 SSTGSDIMVDSPLATLADEVASLAKOGL 758
Db 1799 SSSGSINLLESPQLATLAEDVTAALAKOGL 1828

RESULT 4
Q28869
ID Q28869 PRELIMINARY; PRT; 323 AA.
AC Q28869;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein 2 (MAP2) (MAP-2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RC SEQUENCE FROM N.A., AND CHARACTERIZATION OF TUBULIN REPEATS.
RC TISSUE=BRAIN;
RX MEDLINE=95034751; PubMed=7947727;
RA Coffey R.L., Joly J.C., Cain B.D., Purich D.L.;
RT "Exploring the microtubule-binding region of bovine microtubule-
associated protein-2 (MAP-2): cDNA sequencing, bacterial expression,
and site-directed mutagenesis.";
RL Biochemistry 33:13199-13207(1994).
RN [2]
RN SEQUENCE OF 121-139.
RX MEDLINE=90037224; PubMed=2808529;
RA Joly J.C., Flynn G., Purich D.L.;
RT "The microtubule-binding fragment of microtubule-associated protein-2:
location of the protease-accessible site and identification of an
assembly-promoting peptide.";
RL J. Cell Biol. 109:2289-2294(1989).
RN [3]
RN SEQUENCE OF 18-36 AND 227-258.
RX MEDLINE=92011652; PubMed=1717454;
RA Dingus J., Obar R.A., Hyams J.S., Goedert M., Vallee R.B.;
RT "Use of a heat-stable microtubule-associated protein class-specific
antibody to investigate the mechanism of microtubule binding.";
RL J. Biol. Chem. 266:18854-18860(1991).
RN [4]
RN PHOSPHORYLATION.
RX MEDLINE=98026080; PubMed=9376363;
RA Itoh T.J., Hisanaga S., Hosoi T., Kishimoto T., Hotani H.;
RT "Phosphorylation states of microtubule-associated protein 2 (MAP2)
determine the regulatory role of MAP2 in microtubule dynamics.";
RL Biochemistry 36:12574-12582(1997).
CC -!- FUNCTION: HEAT STABLE PROTEIN WHICH PROMOTES MICROTUBULE ASSEMBLY
```


AND STABILITY, AND MIGHT BE INVOLVED IN THE ESTABLISHMENT AND MAINTENANCE OF NEURONAL POLARITY. MAY PLAY A ROLE IN DENDRITIC POLARITY, OUTGROWTH, BRANCHING, PLASTICITY, SPACING, AND STABILITY OF MICROTUBULES DURING DIFFERENT STAGES OF NEURONAL DEVELOPMENT. THE VARYING LENGTH OF THE PROJECTION ARM SERVES AS A SPACER BETWEEN MICROTUBULES.

-|- ALTERNATIVE PRODUCTS: AS IN OTHER MAMMALS, SEVERAL ISOFORMS INCLUDING ONE POSSESSING AN ADDITIONAL TAU/MAP REPEAT, MIGHT BE PRODUCED BY DEVELOPMENTALLY CONTROLLED ALTERNATIVE SPLICING.

-|- DOMAIN: THE REPEATED DOMAIN BINDS TO TUBULIN. THE SECOND REPEAT PLAYS A DOMINANT ROLE IN MICROTUBULE BINDING.

-|- PTM: PHOSPHORYLATION AT VARIOUS SITES IN THE N-TERMINAL BY A CAMP-DEPENDENT PROTEIN KINASE (PKA), IN THE PROLINE-RICH C-TERMINAL DOMAIN BY PROLINE-DIRECTED PROTEIN KINASES (CDC2, GSK3) AND OF SERINE RESIDUES IN K-X-G-S MOTIFS BY MAP/MICROTUBULE AFFINITY-REGULATING KINASE (MARK). IN VITRO PHOSPHORYLATION BY PKA AFFECTS THE MICROTUBULE-ASSEMBLING ACTIVITY, BY CDC2 AFFECTS BOTH THE MICROTUBULE-ASSEMBLING AND STABILIZING ACTIVITY OF MAP2.

-|- PHOSPHORYLATION BY MARK DECREASES TUBULIN-BINDING AND MICROTUBULES ASSEMBLY, ALLOWING THE REARRANGEMENT OF THE MICROTUBULE NETWORK.

-|- PTM: THE N-TERMINAL IS BLOCKED.

-|- MISCELLANEOUS: MUTAGENESIS EXPERIMENTS WERE DONE IN REF 1 IN WHICH THE FIRST AND THIRD REPEATS WERE MODIFIED TO RESEMBLE THE SECOND REPEAT WHICH LEAD TO SOMEWHAT HIGHER AFFINITY MICROTUBULE BINDING.

-|- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS, FROM WHICH ONE IS APPROXIMATE.

EMBL: S74025; AAB32526.1; ALT_SEQ.

InterPro: IPR001084; Tubulin_Tau.

Pfam: PF00418; tubulin-binding; 3.

PROSITE: PS00229; TAU_MAP; 2.

Microtubules; Repeat; Alternative splicing; Cytoskeleton;

Phosphorylation.

NON_TER 1

REPEAT 156 187

REPEAT 198 218

REPEAT 219 230

REPEAT 251 282

MOD_RES 37

MOD_RES 109

MOD_RES 112

MOD_RES 115

MOD_RES 175

MOD_RES 206

MOD_RES 238

MOD_RES 291

MOD_RES 295

MOD_RES 298

SEQUENCE 323 AA; 34131 MW; 1C0025E15CE835F4 CRC64;

Query Match 18.0%; Score 715; DB 6; Length 323;

Best Local Similarity 47.6%; Pred. No. 6.3e-33;

Matches 161; Conservative 47; Mismatches 72; Indels 58; Gaps 9;

433 PEPPSS-PKHVSSTVSRGSSGAKEMKLGADGKTKIATPRGAAPP- - - - -GQKQANATR 487

32 PEKRSLLPRSSLLPPRGVSGDRDENSEFSLNSISSARRITRSEPIRRACKSGSTPT- 90

488 IPAKTPPAPKTPPSGCEPPKSGDRSGYSS- - - - -PGSGCTGCSRRTPSLPTTP- - - - -TREP 540

91 TPGSTAITPGTTPS- - - - -YSSRTPTGTPCTP- - - - -SYPRTHPTGTPKSAIILVPS 137

541 KKVAVVTRTPKPSKSSAKSLQAPVMPDLKNVKSIGSTENLKHOPGGKGVQIINKLID 600

138 KKVAVIIRTPPKPATPK-QLRINLPDLKNVKSIGSTDNIRKIQPKGGQVQIVTRKID 196

601 LSNVQSKGSKDNIIKHVPGGSGVQIVKVDLSKVTSKGSLGNIHHKPGGQGVKSEK 660

197 LSH- - - - -VTSKGSLNIHRHPPGGGRVKIESVK 225

661 LDFKDRVQSKIGSLDNITHVPGGNGKNIETHKLTFRENAKAKYTDHGAIEIVKSPVVGDT 720

226 LDFKEKAQAKVGSILDNAHVPGGNGVKNIDSKLNFREHAKARVDHGAIEITQSPGRSSVA 285

721 SPRHLNVSSTGSDMDVDSPLQATLADDEVASLAKOGL 758

286 SPRLSNVSSSGSINLLESPLQATLADDEVTAALAKOGL 323

RESULT 5

Q95KDL PRELIMINARY; PRT; 801 AA.

AC Q95KDL; 2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Hypothetical 85.4 kDa protein.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Eutelesostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Macaca.

NCBI_TaxID=9541;

SEQUENCE FROM N.A.

RC TISSUE=MEDULLA OBLOGATA;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.;

RT "Isolation of full-length cDNA clones from macaque brain cDNA

libraries."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB062954; BAB60738.1; -

DR InterPro; IPR001084; Tubulin_Tau.

DR Pfam; PF00418; tubulin-binding; 4.

DR PROSITE; PS00229; TAU_MAP; UNKNOWN_4.

KW Hypothetical protein.

SQ SEQUENCE 801 AA; 85445 MW; 82C4193CB4DA4C6F CRC64;

Query Match 17.8%; Score 704; DB 6; Length 801;

Best Local Similarity 29.4%; Pred. No. 7.8e-32;

Matches 237; Conservative 99; Mismatches 271; Indels 198; Gaps 33;

28 GYTHQDQEGDTAGLKESPLQPTDSEEPGSETSDAKSTPTAEDVTAFLVDEGAPGK 87

92 GNSVHQ-KTSEFHLGLIEG- - - - -DKNKTIPVQKVA- - - - -GKNGL 128

88 QAAQAPHTPEGTABEAGIGDTPSLEDAAGHVTOEPESGVVQEGFLREPPGLSH 147

129 ETKSQSLDF- - - - -LGAADVPTRYAKEQETSVMNPNFHPVAQSPGLSGREATPG- - - - -E 178

148 QLMGMPGAPLLPPEPREATQPSGTGPDTEGHRHAPPELLKQLLDLHQQEGPPLKAG 207

179 MKNIIIPGCPVI- - - - -GVVNNSEQLKCESPLVSVAHAP- - - - -PTTIPITMVF 230

208 GKRPGSKEEV-DEDRVDVE-SSPDQSPPSKASPAQD- - - - -G 243

231 TOEHLNASSHIRDHDELEKLSSTEEEAVALNQAPQKTVRRALSECSHLVSVPVNLAD 290

244 RPPOTAAREATSIPTFAEAGIPLVDFLSKVSTEIPASEPDGSPVSGRA- - - - - 292

291 KYPELPAREE- - - - -PSSGLLPPP- - - - -SSPSSPMP-GKLGAAPAMKRSMTVGEQATSY 340

293 KGQAPALEFTFHVBITNV- - - - -QKQAH- - - - -SEHLGRAAPFAPG- - - - -E 333

341 KLSPGKLPVISTKEIPPFICEBPVAKKRELAHFNSNSSSGKKEL-GTAGLHLHSLKEQ 399

334 GPEARGPSLG-EDPKADLPES- - - - -EKQPA-AAPRGKPVSRVPQLKARVMSKDG 385

400 IPEGSSKEKQEDISEPRIDSCSVQCGKEQPGQTALACKKEIV- - - - - 445


```
QY 386 TGSDDKAKKTSTRSS-----AKTLKNRPCLSL-----PKLPTPGSSDPL 423
Db 446 -----TATQSTSFCEEPDRGIAPEGRPAVSGTNDITTPPNKELPPSPK 496
QY 424 IQPSSAVCPPEPPSP-KHVSSVTSRTSGGAKEMKLGADGKTKIATPRGAAPQKQ 482
Db 497 TKPIADAKAKRASPSPKASPARSGS-----KSTQIVAKATTASSVASTGPSR 548
QY 483 ANATRIAPAK-----PPAPKTPPSGCEPPKSGDRSGYSPGSPGTPGSRRT 530
Db 549 SPSTLLPKKPTADLSRPKSTSTSMKKTTLTSGTAPAGVAPTRVKP-TP-MPSRPSTP 606
QY 531 SLPTPTPREPKKVVAVRTPPKSPSSAKSRLQ--TAPVMPDLKNVSKIGSTENLKHQP 588
Db 607 SIDKKPT-----SAPSSSTPRLRLATNASAPDLKNVRSKVGSTENLKHQP 654
QY 589 GKGVOIINKLDLSNVQSCGSKNDTKHVPVGGSGVOIVYKVDLSKVTSCGSLGNIHKK 648
Db 655 GGRVQIVSKVSYHSIQSCGSKNDIKHVPVGGNVQIQNKVDISKVSKCGSKANIKH 714
QY 649 PGGGVQEVKSEKLDKDRVQSKIGSLDNIHVPGGNGKKTIEHKLTFRENAKAKTDHGA 708
Db 715 PGGGDVKEISQKLNFEKAKQKVGSLDNGVHLPAGGAVKIETRYLTFRANARARTDHG 774
QY 709 IYKSPVWSG--DTSRPHLSNVSS 731
Db 775 IVSRPPHFGPGNSGRVLGLSLRT 799

RESULT 6
P97749
ID P97749 PRELIMINARY; PRT; 198 AA.
AC P97749;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96263369; PubMed=8847742;
RA Vouyiouklis D.A., Brophy P.J.;
RT "Microtubule-associated proteins in developing oligodendrocytes:
RT transient expression of a MAP2c isoform in oligodendrocyte
RT precursors.;"
RL J. Neurosci. Res. 42:803-817(1995).
CC -!- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
DR EMBL; S82901; AAB46832.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 4.
DR PROSITE; PS00229; TAU_MAP; 3.
KW Microtubules; Repeat.
FT NON_TER 1
SQ SEQUENCE 198 AA; 21397 MW; 534719226E846078 CRC64;

Query Match
Best Local Similarity 17.4%; Score 690.5; DB 11; Length 198;
Matches 130; Conservative 35; Mismatches 30; Indels 1; Gaps 1;

QY 541 KKVAVVTRPPKSPSSAKSRLQATPVMPDLKNVSKIGSTENLKHQPGGKGVQIINKLD 600
Db 4 KKVAIIRTPPKSPATPK-QLRLLNQPLDLKNVSKIGSTDNIKYQPGQVRILNKKID 62
QY 601 LSNVQSCGSKDNIKHVPVGGSGVOIVYKVDLSKVTSCGSLGNIHKKPQGGQEVKSEK 660
Db 63 FSKVQSRGSKDNIKHSAGGNGVQIVTKKIDLSHVTSCGSLKNIRHRPQGGRVKIESVK 122
QY 661 LDFKDRVQSKIGSLDNIHVPGGNGKKTIEHKLTFRENAKAKTDHGAIVYKSPVWSGT 720
Db 720 IYKSPVWSG--DTSRPHLSNVSS 731
```

```
Db 123 LDFKEKAQKVGSLDNAHHVPGGNGVKIDSKLNPREHAKARVDHGAEIITQSPSRSSVA 182
QY 721 SPRHLSNVSSGSDM 736
Db 183 SPRRLSNVSSGSLN 198

RESULT 7
Q64710
ID Q64710 PRELIMINARY; PRT; 381 AA.
AC Q64710;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Microtubule-associated protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95046346; PubMed=7957935;
RA Langkopf A., Guilleminot J., Nunez J.;
RT "Two novel HMW MAP2 variants with four microtubule-binding repeats and
RT different projection domains.;"
RL FEBS Lett. 354:259-262(1994).
CC -!- SIMILARITY: CONTAINS 3 TAU/MAP REPEATS.
DR EMBL; S74265; AAB32559.1; -.
DR InterPro; IPR001084; Tubulin_Tau.
DR Pfam; PF00418; tubulin-binding; 3.
DR PROSITE; PS00229; TAU_MAP; 2.
KW Alternative splicing; Microtubules; Repeat.
FT NON_TER 1
FT VARSPLIC 83 127 MISSING (IN SHORT FORM).
SQ SEQUENCE 381 AA; 41794 MW; ADF65EBF9D35F8D0 CRC64;

Query Match
Best Local Similarity 14.5%; Score 576; DB 11; Length 381;
Matches 146; Conservative 57; Mismatches 97; Indels 78; Gaps 14;

QY 336 EARGPSLGEDTFKADLPPESEKQPAAPRGKPVSRVQPKARVMSKDGTSDDKAKT 395
Db 35 EARRPSLEKHKE-----KPFKTCRG----RISTPERKVAKEPSTVSRDEVRKK 81
QY 396 STRSSAKTLKNRCLSPKLPPTPGSSDPLIQFSPA--VCPEPP---SPKHSVSVTSRT- 449
Db 82 AVYKKAELAK-----ESEVQAHSPSRKLLKPAIKYTRPHTLSCVKKRKT 126
QY 450 -----GSSGAKEMKLGADGKTKIATPRGAAP-----PGQKG-----QANATRI 490
Db 127 ATSGESAQAPSAFQAKDKVTDGITKSPKRSLPRPSSILPPRRGVSGDRENSFSLNS 186
QY 491 KTPPAPKTPPSSGEPKSGDRSGYSPGSPG-----TPGS-----RSRTPSLPTPT--REP 540
Db 187 SISSARRTTRS--EPIRRACKSGTSTTPTGTAITGTPPSYSSRTPGTPGTPSYRTP 244
QY 541 -----KKVAVVTRPPKSPSSAKSRLQATPVMPDLKNVSKIGSTENLKHQP 588
Db 245 GTPKSGILVPESEKVAIIRTPPKSPATPK-QLRLLNQPLDLKNVSKIGSTDNIKYQK 303
QY 589 GKGVOIINKLDLSNVQSCGSKDNIKHVPVGGSGVOIVYKVDLSKVTSCGSLGNIHKK 648
Db 304 GQVRIILNKIDFSQVSRGSKDNIKHSAGGNGVQIVTKKIDLSHVTSCGSLKNIRH 363
QY 649 PGGGVQEVKSEKLDKDR 666
Db 364 PGGGRVKIESVKLDFKEK 381

RESULT 8
Q9DBX9
ID Q9DBX9 PRELIMINARY; PRT; 414 AA.
```


AC Q9DBX9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Microtubule-associated protein.
 GN MTAP4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LUNG;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Wyshaw-Borls A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -/- SIMILARITY: CONTAINS 4 TAU/MAP REPEATS.
 DR EMBL; AK004694; BAB23480.1; -;
 DR MGD; MGI:97178; Mtap4.
 DR InterPro; IPR002965; P-rich_extensin.
 DR Pfam; PF00418; tubulin-binding; 4.
 DR PRINTS; PR01217; PRICHEXTENSIN.
 DR PROSITE; PS00229; TAU_MAP; 3.
 KW Microtubules; Repeat.
 SQ SEQUENCE 414 AA; 41784 MW; DFBD695FE8FF9D40 CRC64;
 Query Match 14.3%; Score 567.5; DB 11; Length 414;
 Best Local Similarity 36.7%; Pred. No. 1.7e-24;
 Matches 143; Conservative 50; Mismatches 100; Indels 97; Gaps 12;
 QY 351 LPEPSEKQAAAPRGKPVSRVP--QLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKLRP 408
 DB 8 VPAAPHKRAAATATARPSTLPARDVKPKPITEAKVA---EKRTSPSKPSSAPALKPGP 63
 QY 409 CLSPKLTPTGSSDPLIO--PS--SPA--VCEPPSS-----PKHSSVTSRTGSSAKEM 457
 DB 64 KTTPTVSKATSPSTLVSTGSSRSRPAATTLKRPSTIKTEGKPADVVRMTAKSASADLSRS 123
 QY 458 KLKADGCKTIATPRGAAPQKQOANATRIAPAKTPPAPKTPPSSGPPKSGDRSGYSSP 517
 DB 124 KTTASASVKRNTPTGAAPP---AGMTSTRV-----KPSAAPSRSAGLSVDKK----- 169
 QY 518 GSPGTPGSRSTRPTSLPTTPPREKVAVVRTPPKSPSSAKSRQLQTAPVPMPLKNYKSKI 577
 DB 170 -----PT-----STKPSSSAPRVSRLAT--VSAPDLKSVRSKV 201
 QY 578 GSTENLKHQPGG-----GKQVQINKKL 599
 DB 202 GSTENIKHOPGGGRAKVEKTEATATTAGKPEPNNAVTKAAGSIASAKPPAGKQVIVSKKV 261
 QY 600 DLSNVQSKCGSKDNKIVPGGSGVQIVYKPVLDLSKVTSCGSLGNTLHKPGGQVEVKSE 659
 DB 262 SYSHIQKCGSKDNKIVPGGSGVQIONKVDISKVSSKCGSKANIKHKPGGQVIESQ 321

QY 660 KLDFKDRVQSKIGSLDNITHVPGGNGKKIE 689
 DB 322 KLNFEKAQAQKVGSLDNVGHLPAGGAVKTE 351
 RESULT 9
 Q8T5L5
 ID Q8T5L5 PRELIMINARY; PRT; 838 AA.
 AC Q8T5L5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative microtubule binding protein.
 GN 4F11.1.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 OC Anopheles.
 OC NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Thomasova D., Ton L.Q., Collins F.H., Kafatos F.C.;
 RT "Sequencing and analysis of Pen1 region from Anopheles gambiae
 RT Chromosome 2R.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ438610; CAD27473.1; -;
 SQ SEQUENCE 838 AA; 88130 MW; E4A2EBAEAB04A26 CRC64;
 Query Match 12.2%; Score 485; DB 5; Length 838;
 Best Local Similarity 24.9%; Pred. No. 1.8e-19;
 Matches 208; Conservative 76; Mismatches 269; Indels 282; Gaps 35;
 QY 47 PLOTPTEDGSEPGSETSDAKSTPTAEDVTAPLV-----DEG----- 83
 DB 106 PSQPTTTFAPRAEVKFSVPLKTPVPRLLPQQOQHQRDTGPALFPAPISHRPP 165
 QY 84 -----APGKQAAQHPTEIPEG--TTAEAEAGIG--DTPSLEDEAAGHVTPQESKV 131
 DB 166 PIAHQQAAPFAMPARPMPGPMMPGQMPMPGNGVGPRTGT-----QPQPPRPGM 217
 QY 132 VOEGFLRPPGGLSHQLMGMGAPALLPEG-----PREATROPSTGTPEDTEGHRAP 185
 DB 218 Y-----PQPPGVP-----MPMRPQMPGAVPMQPMQPRPPSAOG----- 253
 QY 186 ELLKHOLLCDLHQEGPLKAGGKERP-----GSKEEVEDRDVDESSPOSPSPSKASPA 240
 DB 254 -----MQRPMPGQPPPIRPNMGPPRQI-----SPQSNLSGGMP 292
 QY 241 QDCRPPTAAREATSIPIGFPAGAIPLPVDFLSKVSTIPIASEPDGSPVGRAKGODAPLE 300
 DB 293 GMVGPFRPMPMQGAGPGPGQMRP---NFYNRPMGDPQTSRP---PSGNDNMGGGPPPS 347
 QY 301 FTFHVEITPNVQKQAEHHLGRAAFPCA-----PGEPEAR-----GPSIGEDTK 347
 DB 348 -----SATPSVDDDE----DVYIGRLPADNSALNSPNPARAPPNFTWPGPGGIGEREK 399
 QY 348 EADLPESKQPAAPRGKPVSRVPQLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKN- 406
 DB 400 S-----NPSRPPSVAGSYGKP-----NDHELDSSGGRPLHALKDF 435
 QY 407 -----RPLSP-KLPTPGSSDPLTOPSSPA----- 430
 DB 436 INKEPRPCQSPQSPSGSQSL-----SPANTDENFSYRPGCAKPNSGQOQOQOQOQOQ 491
 QY 431 -VCEP- -PSSPKHVSVTSRTGSSGAKEMKLGADG-KTKIATPRGAAPGQK- - - 481
 DB 492 KLOPPPGGRPNAPNPSAVTPGGR-----AEGDKVTFQIPNGGGGGGGGGGREG 541
 QY 482 --QANATRIAPAKTPPAPKTPPSSGPPKSGDR-----SGYSPGSPGTPGS 525
 DB 542 SQEWSRSRPPQOHSMLRTGPKSLAPDHEGDNDSGVDEYTBQKDRPNALASPSP--LKS 599


```
Db 138 -----KPSPP--IPDAPTMEDIAPRELESJNFSETSGTSDQADRI 176
QY 458 KLGADGKTIATPRGAAPPKGOANATRIPAKTPAPKTPPSG-EPPKSGDRSGYSS 516
Db 177 MQNENERVE-----EKKQMS-----PTPSOPQHKTPORGIRPPTAILRQPKPI 221
QY 517 PGSPGTPGSRSTPSLPTPTREPKKAVVVRTPPKSPSSAKSRLOQTAPVPMPLDK----- 571
Db 222 PA-----SLPRATATPSSQRAISTPRQTAS-----TAPSPRISKMSRER 262
QY 572 -----NVKSGIGSTENLKHQPGGKVOIINKKLDLSNV 604
Db 263 SDVQKSTSTRIDNVRMTPKVNAKFNKSVKSVTN--HKAGGNVEIFSEK-RLYNA 319
QY 605 QSKCGSKDNKHPGGGSGVQIVYKVPVDSL-----KVT----- 636
Db 320 QSKVGLKNATHVAGGNGVQIENRKLDFAASPKVGSKTNYOPAKSDVKIVSEKLTWQAK 379
QY 637 SKCGLGNIHHKPGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGKNIETHKLTFR 696
Db 380 SKVGSMDNAHAKPAGNVQILSQKLNWK--AESKVGSKDNMNRPGGNGVQIFDEKIRY- 436
QY 697 ENAKAKTDHGAELVYKSPVVGDTSPRHLS-NVSS 730
Db 437 -----VSTDSSRNHSTLDISS 452
```

Search completed: January 28, 2003, 14:18:26
Job time : 46.3037 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 42.2376 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQEFVEMDHAGTYGL.....SPQLATLADEVSASLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116.5	53.4	441	15 AAR58810	Human tau protein.
2	2116.5	53.4	441	21 AAY81386	Human paired helic
3	2116.5	53.4	441	23 ABB0398	Human tau protein.
4	2098.5	52.9	441	23 AAE14577	Human Cysless htai
5	2083.5	52.5	441	18 AAW34856	Human tau protein.
6	2083.5	52.5	441	23 ABB05642	Mutant human tau p
7	2077.5	52.4	441	17 AAW05282	Human tau protein.
8	1837.5	46.3	390	17 AAW05283	Truncated human ta
9	1590	40.1	352	19 AAY20248	Human microtubule
10	1586	40.0	351	21 AAY15200	Human Tau protein.

11	1574	39.7	352	10 AAR91294	Paired helical fil
12	1574	39.7	352	14 AAR32708	Human tau-protein.
13	1350.5	34.1	364	23 ABB57300	Mouse ischaemic co
14	1329	33.5	314	19 AAY20956	Human microtubule
15	812.5	20.5	1831	18 AAW23329	Microtubule-associ
16	809.5	20.4	1824	19 AAY20300	Human microtubule
17	807	20.4	1828	23 ABB57249	Mouse ischaemic co
18	801.5	20.2	1717	18 AAW23331	Neuroblastoma indi
19	733	18.5	140	17 AAW05285	Human tau protein
20	657	16.6	125	22 AAB00720	Human htau40 pepti
21	621.5	15.7	1125	21 AAB22934	Mouse microtubule-
22	621.5	15.7	1125	21 AAY79637	Mouse microtubule-
23	621.5	15.7	1610	21 AAB22870	Microtubule associ
24	621.5	15.7	1610	21 AAY79648	EFYP-DEVD-NAP4-EBF
25	593	15.0	112	16 AAR76937	Bifunctional caspa
26	563	14.2	106	17 AAR92516	PHF-tau (143-254)
27	551	13.9	140	17 AAW05286	Microtubule-associ
28	501	12.6	95	17 AAW05284	Fragment of microt
29	471	11.9	700	22 ABB61671	Human tau protein
30	470.5	11.9	94	22 AAB00729	Drosophila melanog
31	462.5	11.7	94	22 AAB00732	Human TAU derived
32	456	11.5	91	22 AAB00731	Human TAU derived
33	455.5	11.5	94	22 AAB00730	Human TAU derived
34	419	10.6	78	22 AAB5640	Tau peptide region
35	364	9.2	67	15 AAR59837	Sequence of human
36	300	7.6	1466	22 AAE02537	Porcine alpha(III)
37	298.5	7.5	1466	22 AAE02533	Bovine alpha(III)
38	298.5	7.5	1466	22 AAE02534	Bovine alpha(III)
39	287.5	7.3	99	21 AAG03156	Human secreted pro
40	285.5	7.2	1806	23 AAG04266	Human endometrial
41	283.5	7.2	1466	22 ABB50291	Collagen type III
42	283.5	7.2	1466	23 ABB50747	Human tumour Endot
43	283.5	7.2	1469	22 ABB51591	Novel human diagno
44	283.5	7.2	1745	22 AAG77793	Human pro-alpha-3(
45	283.5	7.2	1745	23 ABB97234	Novel human protei

ALIGNMENTS

RESULT 1

AAR58810
ID AAR58810 standard; protein; 441 AA.

AC AAR58810;

XX 27-MAR-1995 (first entry)

DT Human tau protein.

DE

XX

XX Tau; cerebrospinal fluid; immunoassay; antibody; detection;

XX diagnosis; central nervous system; CNS; cytopathies; cytopathy;

XX Alzheimer's disease.

OS Homo sapiens.

XX WO9418560-A.

XX 18-AUG-1994.

XX 10-FEB-1994; 94WO-JP00196.

XX 12-FEB-1993; 93JP-0046133.

XX (TEIJ) TEIJIN LTD.

XX Eguchi H, Hosoda K, Kobayashi S, Kubota T, Mori H;

XX Nakamoto T;

XX WPI; 1994-279910/34.

XX Sandwich immunoassay of tau protein in cerebrospinal fluid - for

XX diagnosis of Alzheimer's disease and other CNS cytopathies

PT


```

XX PS Claim 1; Page 16-18; 36pp; Japanese.
XX CC Detection of the human tau protein (or fragments of it) in samples
XX CC of cerebrospinal fluid enables the diagnosis of central nervous
XX CC system cytopathies such as Alzheimer's disease. Detection is
XX CC performed using labelled antibodies which recognise sites within the
XX CC region defined by the amino acid residues 251-441. The antibodies
XX CC are preferably polyclonal.
XX SQ Sequence 441 AA;

Query Match 53.4%; Score 2116.5; DB 15; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

Qy 1 MAEPRQFEVMEHAGTYGLGDRKQGGYTHQDQGGTDAGLKESPLQTPTEGSEEPG 60
Db 1 MAEPRQFEVMEHAGTYGLGDRKQGGYTHQDQGGTDAGLKESPLQTPTEGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGKVVOEGFLREPFGPGLSHQMLSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQ----- 124
Qy 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDSSPDSPSKASPA 240
Db 125 ----- 124
Qy 241 QDGRPPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Db 125 ----- 124
Qy 301 FTFHVEITPNVQKEQAHSEHGLRAAFGAPGEGPEARGPSLGEDTKEADLPPESEKQPA 360
Db 125 ----- 124
Qy 361 AAPRGKPVSRYPQLKARVSKSDGTGDDKAKTSTRSSAKTLKRNPCLSPKLPTPGSS 420
Db 125 -----ARVSKSDGTGDDKKA----- 142
Qy 421 DPLQSPSPAVCPPEPPSPKHSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPGQK 480
Db 143 -----KGADGKTKIATPRGAAPGQK 163
Qy 481 GOANATRIPTAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 164 QGANATRIPTAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 223
Qy 541 KKVAVRTPPKSPSAKSRLOTAPVMPDLKNVSKIGSTENLKHOPGGKVQIINKLD 600
Db 224 KKVAVRTPPKSPSAKSRLOTAPVMPDLKNVSKIGSTENLKHOPGGKVQIINKLD 283
Qy 601 LSNVQSKGSKDNKIHVPGGSGVQIVYKPVDSLKVTSKCGSLGNHHPGGGVQEVKSEK 660
Db 284 LSNVQSKGSKDNKIHVPGGSGVQIVYKPVDSLKVTSKCGSLGNHHPGGGVQEVKSEK 343
Qy 661 LDFKDRVQSKIGSLDNTHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 720
Db 344 LDFKDRVQSKIGSLDNTHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 403
Qy 721 SPRHLNSVSTGSDIMVDSPOLATLADRVSVASLAKOGL 758
Db 404 SPRHLNSVSTGSDIMVDSPOLATLADRVSVASLAKOGL 441

```

RESULT 2
AAV81386
ID AAV81386 standard; protein; 441 AA.
XX

```

AC AAV81386;
XX DT 19-JUN-2000 (first entry)
XX DE Human paired helical filament phosphorylated tau protein.
XX KW Phosphorylated tau protein; human; paired helical filament;
XX KW polyclonal antibody; Alzheimer's disease; cerebrospinal fluid; CSF;
XX KW diagnosis; detection.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Modified-site 199
FT Modified-site /note= "Phosphorylated"
FT Modified-site 231
FT Modified-site /note= "Phosphorylated"
FT Modified-site 235
FT Modified-site /note= "Phosphorylated"
XX JP20000034300-A.
XX PN 02-FEB-2000.
XX PD
XX PF 17-JUL-1998; 98JP-0204040.
XX PR 17-JUL-1998; 98JP-0204040.
XX PA (MITU ) MITSUBISHI CHEM CORP.
XX XX
XX DR WPI; 2000-285529/25.
XX PT Anti-phosphated tau protein antibody - for the detection of Alzheimer
XX PT disease
XX PS Claim 1; Page 8-9; 12pp; Japanese.
XX CC The invention relates to an antibody against a phosphorylated tau
XX CC protein (AAV81386), which is a component of the paired helical filament
XX CC found in the plaques associated with Alzheimer's disease. A
XX CC phosphorylated tau protein fragment selected from peptides
XX CC AAV81387-Y81390 is conjugated to keyhole limpet haemocyanin (KLH), and
XX CC used to raise polyclonal antibodies in a rabbit. The antibodies of the
XX CC invention are specific for phosphorylated tau protein and may be used to
XX CC detect phosphorylated tau protein in the cerebrospinal fluid (CSF) of a
XX CC patient suspected of having Alzheimer's disease. Use of the antibodies of
XX CC the invention provides specific diagnosis of Alzheimer's disease. The
XX CC present sequence represents phosphorylated tau protein.
XX SQ Sequence 441 AA;

Query Match 53.4%; Score 2116.5; DB 21; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

Qy 1 MAEPRQFEVMEHAGTYGLGDRKQGGYTHQDQGGTDAGLKESPLQTPTEGSEEPG 60
Db 1 MAEPRQFEVMEHAGTYGLGDRKQGGYTHQDQGGTDAGLKESPLQTPTEGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGKVVOEGFLREPFGPGLSHQMLSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQ----- 124
Qy 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDSSPDSPSKASPA 240
Db 125 ----- 124
Qy 241 QDGRPPQTAAREATSIPIGFAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300

```


Db 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHLGRAAPGAPGEGPEARGLSGEDTKADLPPEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARWVSKSDGTGSDDKRAKTSRSTSSAKTLKRNPCLSPLKLPSS 420
Db 125 ----- 142
QY 421 DPLIQSSPAVCEPSSPKHVSSTRTGSSGAKEMKLKADGKTKIATPRGAAPPQOK 480
Db 143 ----- 163
QY 481 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTSRSTPSLPTPTREP 540
Db 164 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTSRSTPSLPTPTREP 223
QY 541 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKLKD 600
Db 224 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKLKD 283
QY 601 LSNVQSKCGSKDNKIHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 660
Db 284 LSNVQSKCGSKDNKIHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 343
QY 661 LDFKRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 720
Db 344 LDFKRVQSKIGSLDNITHVPGGKNIETHKLTFRENAKAKTDHGAIEIVKSPVSGDT 403
QY 721 SPRHLSNVSTGSDMWDSQPLATLADEVASLAKQGL 758
Db 404 SPRHLSNVSTGSDMWDSQPLATLADEVASLAKQGL 441

RESULT 3
ABB09398
ID ABB09398 standard; peptide: 441 AA.
XX ABB09398;
AC ABB09398;
DT 01-JUL-2002 (first entry)
DE Human Tau protein.
XX
KW NF-H; neurodegenerative disorder; immunogenic epitope; neurotropic;
KW neuroprotective; antiparkinsonian; cerebroprotective; vasotrophic;
KW Alzheimer's disease; multi-infarct dementia; Pick's disease;
KW frontotemporal dementia; dementia pugilistica; vascular dementia;
KW Parkinson's disease; Gerstmann-Straussler-Scheinker disease;
KW multiple sclerosis; Amyotrophic lateral sclerosis;
KW transient ischaemic attack; stroke; NF-M; Tau; B-amyloid protein.
XX
OS Homo sapiens.
XX
XX WO200115655-A2.
XX
PD 08-MAR-2001.
XX
XX 27-AUG-2000; 2000WO-IL00509.
XX
PR 31-AUG-1999; 99US-0386347.
PR 27-JUL-2000; 2000US-0221150.
XX
XX (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.
PA Michaelson DW;
XX
XX WPI; 2002-179260/23.
XX
XX Identifying existence, non-existence, type or state of
PT neurodegenerative disorder, by reacting with serum a peptide derived
PT from protein associated with the disorder, to which antibody is

produced at onset of disorder -
Disclosure; Page 116-117; 117pp; English.
XX
XX
XX
CC The invention relates to identifying existence, non-existence, type or
state of neurodegenerative disorder (ND) in individual, involving
immunoreacting with serum derived from an individual, a peptide
representing an epitope derived from endogenous protein to which an
antibody (Ab) is produced in vivo at onset or during progression of ND,
where Ab immunobinds with the peptide and detecting presence, absence or
degree of immunobinding. Activity of the immunogenic peptides of the
invention may be described as, neurotropic, neuroprotective,
antiparkinsonian, cerebroprotective and vasotropic. The peptides of the
invention are useful for identifying existence, non-existence, type or
state of ND in an individual, where ND is associated with progressive
loss of cognitive functions or motoric functions, or progressive loss
of control of motoric functions. Such diseases include; Alzheimer's
disease, Multi-infarct dementia (MID), Pick's disease, frontotemporal
dementia, Gerstmann-Straussler-Scheinker disease, Parkinson's
disease, Amyotrophic lateral sclerosis (ALS), transient ischaemic
attack (TIA) and stroke. In a preferred embodiment of the present
invention, the endogenous protein is selected from the group consisting
of NF-H, NF-M, tau and B-amyloid protein. The current sequence represents
the amino acid sequence of the human Tau protein.
XX
SQ Sequence 441 AA;
Query Match 53.4%; Score 2116.5; DB 23; Length 441;
Best Local Similarity 58.2%; Pred. No. 6.9e-124;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMEDHAGTYGLGDRKQGGYTMHQDEGDTAGLKESPLQTPTEGSEEPG 60
Db 1 MAEPQEFVEMEDHAGTYGLGDRKQGGYTMHQDEGDTAGLKESPLQTPTEGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAOHPHTEPEGTAAEAGTGDPSLEDAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAAOHPHTEPEGTAAEAGTGDPSLEDAAG 120
QY 121 HVTQEPESGKVQVQEGFLRPGPGLSHQLMGPGLPGLPEGPATRGQSGTGPETEG 180
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHQLLDLHQEGPPLKAGGKERPGSKEEVDEDRDVEDSSPQSPSPKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIPOFPAEGAIPLPVDFLSKVSTEIPASEPDGSPVGRAGKQDAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHLGRAAPGAPGEGPEARGLSGEDTKADLPPEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARWVSKSDGTGSDDKRAKTSRSTSSAKTLKRNPCLSPLKLPSS 420
Db 125 ----- 142
QY 421 DPLIQSSPAVCEPSSPKHVSSTRTGSSGAKEMKLKADGKTKIATPRGAAPPQOK 480
Db 143 ----- 163
QY 481 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTSRSTPSLPTPTREP 540
Db 164 GOANATRIPAKTPPAPKTPSSGEPKSGDRSGYSPGSPGTSRSTPSLPTPTREP 223
QY 541 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKLKD 600
Db 224 KKVAVRTPKSPSSAKSRLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKLKD 283
QY 601 LSNVQSKCGSKDNKIHVPGGGVQIYKPVVDSKVTSCGSLGNTHHHPGGQVEVKSEK 660

|||||
Db 284 LSNVQSCGSKDNKIHVPGGSGVQIIVKPVDSLKVTSCGSLGNIIHKKPGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNIITHVPGGSKKIIETHKLTFRNKAATDHGAETIYKSPVVSQGT 720
Db 344 LDFKDRVQSKIGSLDNIITHVPGGSKKIIETHKLTFRNKAATDHGAETIYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKOGL 441

RESULT 4
AAE14577
ID AAE14577 standard; Protein; 441 AA.
XX
AC AAE14577;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human Cysless htau40 mutant protein.
XX
KW Tau; polymerisation; free fatty acid; FFA; filament; brain;
KW Alzheimer's disease; neurodegenerative disease; Pick's disease;
KW progressive supranuclear palsy; corticobasal degeneration; FTDP-17;
KW frontotemporal dementia and Parkinsonism linked to chromosome 17;
KW site-directed mutagenesis; human; htau40; mutant; muten.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200210770-A2.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US41512.
XX
PR 31-JUL-2000; 2000US-221777P.
XX
PA (NEUR-) NEURONAUTICS INC.
XX
PI Gamblin TC, Binder Li;
XX
DR WPI; 2002-206223/26.
XX
PT Identifying reagents for ability to effect tau polymerization,
PT comprises altering oxidative conditions for polymerization and/or the
PT oxidative state of a fatty acid used to stimulate polymerization -
XX
PS Example 2; Page 42-43; 56pp; English.
CC The invention relates to a method of identifying reagents that effect
CC tau polymerisation. The method involves adding a free fatty acid (FFA) to
CC a polymerisation buffer to form a FFA-containing buffer, adding tau
CC protein to the buffer to form a tau polymerisation reaction mixture,
CC adding the reagent to the reaction mixture, incubating it to form tau
CC filaments, and finally detecting the tau filaments. The invention
CC also provides a method for regulating the assembly of tau protein in the
CC brain of a mammal, by administering an inhibitor of FA oxidation.
CC The method of the invention is useful for screening agents which effect,
CC preferably prevent and/or inhibit, tau polymer formation in the brains of
CC patients with Alzheimer's disease and other related neurodegenerative
CC diseases characterised, at least in part, by tau polymer formation
CC e.g. Pick's disease, progressive supranuclear palsy, corticobasal
CC degeneration and frontotemporal dementia and Parkinsonism
CC linked to chromosome 17 (FTDP-17). The improved protocols of the
CC invention for inducing tau polymerisation utilise modifications of the
CC oxidative environment of the polymerisation reaction or modifications of
CC the oxidative state of components of the in vitro polymerisation
CC reaction. The present sequence is human Cysless htau40 double-mutant
CC protein with c291A and c322A substitutions, used in an
CC exemplification of the invention.
XX

SQ Sequence 441 AA;
Query Match 52.9%; Score 2098.5; DB 23; Length 441;
Best Local Similarity 57.9%; Pred. No. 9.2e-123;
Matches 439; Conservative 0; Mismatches 2; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMEDHAGTYGLGDRKDQGGYTHQDQEGDTAGLKESPLQTTEDGSEBP 60
Db 1 MAEPQEFVEMEDHAGTYGLGDRKDQGGYTHQDQEGDTAGLKESPLQTTEDGSEBP 60
QY 61 SETSDAKSTPTAEADVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLDEAAG 120
Db 61 SETSDAKSTPTAEADVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLDEAAG 120
QY 121 HVTQPESGKVQEGFLREPFGPLSHQIMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHLGLDHLQEGPLKAGGKERPGSKKEEVEDEDRDVDESSPDSPSKASPA 240
Db 125 ----- 124
QY 241 QDGRPPQTAAREATSIPGPPAEGAIPLVPVFLSKVSTEIPASEPDGPGSVGRAKGDAPLE 300
Db 125 ----- 124
QY 301 FTFHVEITPNVQKEQAHSEHLGRAAFPGAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPSRVQPKARVYKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLTPTGSS 420
Db 125 -----ARMVSKSDGTGSDDKKA----- 142
QY 421 DPLIQSPSPAVCPPEPPSPKHVSSVTSRTGSGGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 143 -----KGADGKTKIATPRGAAPPQK 163
QY 481 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGPTPGSRSTPSLTPPTREP 540
Db 164 GQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSGPTPGSRSTPSLTPPTREP 223
QY 541 KKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVVTRTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSCGSKDNKIHVPGGSGVQIIVKPVDSLKVTSCGSLGNIIHKKPGGQVEVKSEK 660
Db 284 LSNVQSCGSKDNKIHVPGGSGVQIIVKPVDSLKVTSCGSLGNIIHKKPGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNIITHVPGGSKKIIETHKLTFRNKAATDHGAETIYKSPVVSQGT 720
Db 344 LDFKDRVQSKIGSLDNIITHVPGGSKKIIETHKLTFRNKAATDHGAETIYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSQPLATLADEVASLAKOGL 441

RESULT 5
AAW34856
ID AAW34856 standard; protein; 441 AA.
XX
AC AAW34856;
XX
DT 27-MAR-1998 (first entry)
XX
DE Human tau protein.
XX
KW Antibody; phosphorylated tau protein; paired helical filament;
KW detection; Alzheimer's disease; human.
XX
OS Homo sapiens.

XX PN W09734145-A1.
 XX PD 18-SEP-1997.
 XX PF 13-MAR-1997; 97WO-JP00804.
 XX PR 13-MAR-1996; 96JP-0056090.
 XX PA (MITU) MITSUBISHI CHEM CORP.
 XX PI Inahori K, Ishiguro K, Park J, Sato K, Uchida T;
 XX WP: 1997-470978/43.
 XX
 XX PT Antibody prepared using a partial peptide containing part of
 XX phosphorylated tau protein - used for detecting Alzheimer's disease
 XX
 XX PS Claim 2; Pages 25-27; 48pp; Japanese.
 XX
 CC An antibody, prepared using a partial peptide containing the
 CC phosphorylated residue of the phosphorylated tau protein, e.g. the
 CC present sequence, in a paired helical filament, can be used to
 CC detect Alzheimer's disease, i.e. by detecting phosphorylated tau
 CC protein in brain extracts or tissue fragments.
 XX
 XX SQ Sequence 441 AA;
 Query Match 52.5%; Score 2083.5; DB 18; Length 441;
 Best Local Similarity 57.8%; Pred. No. 7.9e-122;
 Matches 439; Conservative 0; Mismatches 0; Indels 321; Gaps 4;
 QY 1 MAEPQEFVEVMEHDAG--TYGLGDRKQGGYTHQDQEGDTAGLKESPLQPTEDGSEE 58
 Db 1 MAEPQEFVEVMEHDAGQDTYGLGDRKQGGYTH--QEGDTAGLKESPLQPTEDGSEE 58
 QY 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
 Db 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
 QY 119 AGHVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDT 178
 Db 119 AGHVTQ-----
 QY 179 EGGRHAPPELLKOLLGDLHQEGPLKAGGKERPGSKEEVEDRDVDDESSPDSPSKAS 238
 Db 125 -----
 QY 239 PAODGRPPQTAAREATSIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAP 298
 Db 125 -----
 QY 299 LEFTFHEITPNVQKQAHSEHLGRAAPGAPGCEPARGPSLGEDTKREADLPPESEKQ 358
 Db 125 -----
 QY 359 PAAAPRGKPVSRVPLKARMVSKDGTGDDKAKTSTRSSAKTLKNRPLCPKLPPTG 418
 Db 125 -----ARVSKSDGTGSDDKA-----
 QY 419 SSOPLQPSAPVCPPEPPSPKHVSSVTSGTSSGAKEMKLGADGKTKIATPRGAAPP 478
 Db 143 -----KGADGKTKIATPRGAAPP 161
 QY 479 QKQANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSGTGPGRSRTPSLTPPTR 538
 Db 162 QKQANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSGTGPGRSRTPSLTPPTR 221
 QY 539 EPKKVAVVTPPKSPSSAKSRLQTAVPMPDLKNVKSKITSTENLKHQPGGKQVLIINK 598
 Db 222 EPKKVAVVTPPKSPSSAKSRLQTAVPMPDLKNVKSKITSTENLKHQPGGKQVLIINK 281
 QY 599 LDLSNVQSKCGSKDNKIKHVPGGSVQIYKVPDLSKVTSCGSLGNIHHKPGGQVEVKS 658

Db 282 LDLSNVQSKCGSKDNKIKHVPGGSVQIYKVPDLSKVTSCGSLGNIHHKPGGQVEVKS 341
 QY 659 EKLDKDRVQSKIGSLDNITHVPGGNGKNIETHKLTFRENAKAKTDHGAEIYVKSPVVS 718
 Db 342 EKLDKDRVQSKIGSLDNITHVPGGNGKNIETHKLTFRENAKAKTDHGAEIYVKSPVVS 401
 QY 719 DTSPRHLSNVSTGSDMWVDSPLATLADDEVSAKQGL 758
 Db 402 DTSPRHLSNVSTGSDMWVDSPLATLADDEVSAKQGL 441
 RESULT 6
 ABB05642
 ID ABB05642 standard; Protein; 441 AA.
 XX AC ABB05642;
 XX DT 25-APR-2002 (first entry)
 XX DE Mutant human tau protein SEQ ID NO:9.
 XX KW Transgenic mouse; neurofibril; senile dementia; mutant tau protein;
 KW brain; neurodegenerative disease; Alzheimer's disease; human.
 XX OS Homo sapiens.
 OS Synthetic.
 XX PN WO200195709-A1.
 XX PD 20-DEC-2001.
 XX PF 14-JUN-2001; 2001WO-JP05070.
 XX PR 15-JUN-2000; 2000JP-0179723.
 XX PA (RIKE) RIKEN KK.
 XX PI Takashima A;
 XX DR WPI; 2002-097995/13.
 XX PT Transgenic mouse with human mutant tau protein, useful in the
 XX development of treatments for neurodegenerative diseases -
 XX PS Claim 2; Page 28-29; 34pp; Japanese.
 XX
 CC The present invention describes a transgenic mouse transformed by human
 CC mutant tau protein and showing neurofibril changes in the brain due to
 CC senile dementia. Also described is a method for screening for treatments
 CC for neurodegenerative diseases. The transgenic mouse can be used in
 CC developing remedies for neurodegenerative diseases such as Alzheimer's
 CC disease. The present sequence represents a mutant human tau protein
 CC which is used in the exemplification of the present invention.
 XX
 SQ Sequence 441 AA;
 Query Match 52.5%; Score 2083.5; DB 23; Length 441;
 Best Local Similarity 57.8%; Pred. No. 7.9e-122;
 Matches 439; Conservative 0; Mismatches 0; Indels 321; Gaps 4;
 QY 1 MAEPQEFVEVMEHDAG--TYGLGDRKQGGYTHQDQEGDTAGLKESPLQPTEDGSEE 58
 Db 1 MAEPQEFVEVMEHDAGQDTYGLGDRKQGGYTH--QEGDTAGLKESPLQPTEDGSEE 58
 QY 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
 Db 59 PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPGTTAEAGIGDTPSLEDEA 118
 QY 119 AGHVTQEPESGKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDT 178
 Db 119 AGHVTQ-----
 QY 124

QY 179 EGRHAPPELLKHOLLGDLHQEGPPLKAGAGKERPGSKEEVEDRDVDESSPODSPSPSKAS 238
 Db 125 ----- 124
 QY 239 PAQGRPPQTAAREATSIFGFAEGAIPLPVDFLSKVSTEIFASEBDGVSGRAGQDAP 298
 Db 125 ----- 124
 QY 299 LEFTFHEITPNVQKEQAHSEHGLGRAAFPGAGPEGARGPSLGEDTKEADLPPESEKQ 358
 Db 125 ----- 124
 QY 359 PAAAPRGKPVSRVQLKARWVSKDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPTPG 418
 Db 125 ----- 142
 QY 419 SSDPLIQSSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPP 478
 Db 143 ----- 161
 QY 479 OKQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTR 538
 Db 162 OKQANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTR 221
 QY 539 EPKAVAVRTPPKSPSAKSRLQATAPVMPDLKNVSKTIGSTENLKHQPGGKQVQIINK 598
 Db 222 EPKAVAVRTPPKSPSAKSRLQATAPVMPDLKNVSKTIGSTENLKHQPGGKQVQIINK 281
 QY 599 LLSNVQSKCGSKDNKHKVPGGSGVQIVYKPVDSLKVTSCGSLGNIHHKPGGGQVEVKS 658
 Db 282 LLSNVQSKCGSKDNKHKVPGGSGVQIVYKPVDSLKVTSCGSLGNIHHKPGGGQVEVKS 341
 QY 659 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVWSG 718
 Db 342 EKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAIEIVYKSPVWSG 401
 QY 719 DTSRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 758
 Db 402 DTSRHLNSVSTGSDIMVDSPOLATLADEVASLAKQGL 441

RESULT 7

AAW05282

ID AAW05282 standard; Protein; 441 AA.

XX AAW05282;

XX AC AAW05282;

XX DT 20-DEC-1996 (first entry)

XX DE Human tau protein.

XX DE Tau protein; inhibition; modulation; prophylaxis; treatment;

KW Alzheimer's disease; motor neurone disease; Lewy body disease;

KW progressive supranuclear palsy; Pick's disease.

XX OS Homo sapiens.

XX PN WO9630766-A1.

XX PD 03-OCT-1996.

XX PF 25-MAR-1996; 96WO-EF01307.

XX PR 27-MAR-1995; 95GB-0006197.

XX PA (HOFF) HOFMANN LA ROCHE & CO AG F.

XX PI Edwards PC, Harrington CR, Klug A, Roth M, Wischik CM;

XX DR WPI; 1996-455570/45.

XX DR N-PSDB; AAW39591.

XX PT Assay for inhibitors of tau-tau interaction - used for identifying

PT cpds., partic. phenothiazine cpds., for treating pathological
 tau-tau or neuro:filament aggregation

XX Example 2; Page 53-54; 97pp; English.

PS

XX Detecting an agent which modulates or inhibits tau-tau protein
 association comprises contacting two tau proteins, distinct from
 each other yet capable of binding to the other and where one of the
 tau proteins is labelled, in the presence of the agent suspected of
 being capable of modulating or inhibiting tau-tau interaction.
 Agents identified as being modulators or inhibitors of tau-tau
 interaction may be used for the prophylaxis and treatment of
 Alzheimer's disease, motor neurone disease, Lewy body disease,
 Pick's disease or progressive supranuclear palsy.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

RESULT 8
AAW05283
ID AAW05283 standard; Protein; 390 AA.
XX
AC AAW05283;
XX
DT 20-DEC-1996 (first entry)
XX
DE Truncated human tau protein.
XX
KW Tau protein; inhibition; modulation; prophylaxis; treatment;
KW Alzheimer's disease; motor neurone disease; Lewy body disease;
KW progressive supranuclear palsy; Pick's disease.
XX
OS Homo sapiens.
XX
PN W09630766-A1.
XX
PD 03-OCT-1996.
XX
PF 25-MAR-1996; 96W0-EP01307.
XX
PR 27-MAR-1995; 95GB-0006197.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Edwards PC, Harrington CR, Klug A, Roth M, Wischik CM;
XX
DR WPI; 1996-455570/45.
XX
PT Assay for inhibitors of tau-tau interaction - used for identifying
PT cpds., partic. phenothiazine cpds.; for treating pathological
PT tau-tau or neuro:fibrilament aggregation
XX
PS Claim 11; : 97pp; English.
XX
CC Detecting an agent which modulates or inhibits tau-tau protein
CC association comprises contacting two tau proteins, distinct from
CC each other yet capable of binding to the other and where one of the
CC tau proteins is labelled, in the presence of the agent suspected of
CC being capable of modulating or inhibiting tau-tau interaction.
CC Agents identified as being modulators or inhibitors of tau-tau
CC interaction may be used for the prophylaxis and treatment of
CC Alzheimer's disease, motor neurone disease, Lewy body disease,
CC Pick's disease or progressive supranuclear palsy. This sequence of
CC the human tau protein is truncated at amino acid residue 390. The
CC full length protein is given in AAW05282.
XX
SQ Sequence 390 AA;
Query Match 46.3%; Score 1837.5; DB 17; Length 390;
Best Local Similarity 54.6%; Pred. No. 1.4e-106;
Matches 386; Conservative 0; Mismatches 4; Indels 317; Gaps 2;
QY 1 MAEPQEFVNEHAGTYGLDRKQGGYTMHODQGGTDAGLKESPLQPTDGGSEEPG 60
DB 1 MAEPQEFVNEHAGTYGLDRKQGGYTMHODQGGTDAGLKESPLQPTDGGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
DB 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTAAEAGIGDTPSLEDEAAG 120
QY 121 HVTQEPESGKVVQGFLEPREGPPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
DB 121 HVTQ-----
QY 181 GRHAPELLKHOLLGLDHLQEGPLKAGGKRPESKEEVEDRDVDESSPODSPSKASPA 240
DB 125 -----
QY 241 QDGRPPQTAAREATSIQFFAEGAIPLPVDFLSKVSTEIPASEPDGFSVGRAGQDAPLE 300

Db 125 ----- 124
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPCAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
Db 125 ----- 124
QY 361 AAPRGKPVSRVQLKARMVSKSDGTGSDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 125 -----ARVSKSKDGTGSDDKA----- 142
QY 421 DPLIQFSSPAVCPPEPSSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 143 -----KGADGKTKIATPRGAAPPQK 163
QY 481 GQANATRIPAKTPAPKTPSSGEPKSCDRSGYSSPGSGPTPGSRSTRPSLPTPTREP 540
Db 164 GQANATRIPAKTPAPKTPSSGEPKSCDRSGYSSPGSGPTPGSRSTRPSLPTPTREP 223
QY 541 KKVAVVRTPPKSPSSAKSRLQTPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVVRTPPKSLSAKSLQTPVMPDLKNGSKIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVSKCGSKDNIKHVPGGGSQVIVKPVDSLKVTSCGSLGNTHHKGQGVVEKSEK 660
Db 284 LSNVSKCGSKDNIKQVPGGGSQVIVKPVDSLKVTSCGSLGNTHHKGQGVVEKSEK 343
QY 661 LDFKORVQSKIGSLDNITHVPGGNNKKIETHKLTIRENAKAKTDHGA 707
Db 344 LDFKORVQSKIGSLDNITHVPGGNNKKIETHKLTIRENAKAKTDHGA 390
RESULT 9
AAW02048
ID AAW02048 standard; Protein; 352 AA.
XX
AC AAW02048;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human microtubule associated protein Tau wild type fragment.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPP-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Homo sapiens.
XX
PN W09845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98WO-IB00705.
XX
PR 10-APR-1997; 97US-0043163.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX
DR WPI; 1998-609901/51.
DR N-PSDB; AAX75754.
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and

SQ Sequence 351 AA;
 Query Match 40.0%; Score 1586; DB 21; Length 351;
 Best Local Similarity 46.4%; Pred. No. 5.7e-91;
 Matches 351; Conservative 0; Mismatches 0; Indels 406; Gaps 4;
 QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHODQEGDTDAGLKESPLQPTPEDGSEEPG 60
 DB 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHODQEGDTDAGLK----- 44
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAQAHPTEIGTAAEEAGIGDTPSLEDEAAG 120
 DB 45 -----AEEAGIGDTPSLEDEAAG 62
 QY 121 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATRQSGTGPEDETEG 180
 DB 63 HVTQ----- 66
 QY 181 GRHAPELLKHOLLGDLHQBGLPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
 DB 67 ----- 66
 QY 241 QDGRPPQTAAREATSIPGFPAGAIPLPVDLFSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
 DB 67 ----- 66
 QY 301 FTFHVEITPNVQKEQAHSEHGLRAAFPAGGEPGARGPSLGEDTKADLPPESEKQPA 360
 DB 67 ----- 66
 QY 361 AAPRGKPVSRVLPOLKARWVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLTPTGSS 420
 DB 67 -----ARWVSKSDGTGDDKKA----- 84
 QY 421 DPLIQSPSPAVCPPEPPSPKHHVSSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPPQK 480
 DB 85 -----KGADGKTKIATPRGAAPPQK 105
 QY 481 GOANATRIAPKTPAPKTPPPSSGEPKSGDRSGYSSSPGTPGSRSRTPSLPTPTREP 540
 DB 106 GOANATRIAPKTPAPKTPPPSSGEPKSGDRSGYSSSPGTPGSRSRTPSLPTPTREP 165
 QY 541 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 166 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 OS Homo sapiens.

Query Match 39.7%; Score 1574; DB 10; Length 352;
 Best Local Similarity 46.2%; Pred. No. 3.2e-90;
 Matches 350; Conservative 0; Mismatches 2; Indels 406; Gaps 4;
 QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHODQEGDTDAGLKESPLQPTPEDGSEEPG 60
 DB 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHODQEGDTDAGLK----- 44
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAQAHPTEIGTAAEEAGIGDTPSLEDEAAG 120
 DB 45 -----AEEAGIGDTPSLEDEAAG 62
 QY 121 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATRQSGTGPEDETEG 180
 DB 63 HVTQ----- 66
 QY 181 GRHAPELLKHOLLGDLHQBGLPPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
 DB 67 ----- 66
 QY 241 QDGRPPQTAAREATSIPGFPAGAIPLPVDLFSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
 DB 67 ----- 66
 QY 301 FTFHVEITPNVQKEQAHSEHGLRAAFPAGGEPGARGPSLGEDTKADLPPESEKQPA 360
 DB 67 ----- 66
 QY 361 AAPRGKPVSRVLPOLKARWVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLTPTGSS 420
 DB 67 -----ARWVSKSDGTGDDKKA----- 84
 QY 421 DPLIQSPSPAVCPPEPPSPKHHVSSVTSRTGSSGAKEMKLKGADGKTKIATPRGAAPPQK 480
 DB 85 -----KGADGKTKIATPRGAAPPQK 105
 QY 481 GOANATRIAPKTPAPKTPPPSSGEPKSGDRSGYSSSPGTPGSRSRTPSLPTPTREP 540
 DB 106 GOANATRIAPKTPAPKTPPPSSGEPKSGDRSGYSSSPGTPGSRSRTPSLPTPTREP 165
 QY 541 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 166 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 OS Homo sapiens.

RESULT 11
 AAP91294
 ID AAP91294 standard; protein; 352 AA.
 XX AAP91294;
 AC AAP91294;
 DT 18-DEC-1989 (first entry)
 XX Paired helical filament (PHF) core protein.
 DE Paired helical filament (PHF) core protein; Alzheimer's disease;
 XX Paired helical filament (PHF) core protein; Alzheimer's disease;
 KW neurofibrillary tangles.
 XX
 OS

QY 601 LSNVQSKGSDNIKHVPGGSGVQIVYKPVLDLSKVTSCGSLGNIHHKPGGQVEVKSEK 660
 DB 217 -----VQIVYKPVLDLSKVTSCGSLGNIHHKPGGQVEVKSEK 254
 QY 661 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720
 DB 255 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 314
 QY 721 SPRHLSNVSSGSDIMVDSQPLATLADEVSASLAKOGL 758
 DB 315 SPRHLSNVSSGSDIMVDSQPLATLADEVSASLAKOGL 352

RESULT 12
 AAR32708
 ID AAR32708 standard; Protein; 352 AA.
 AC AAR32708;
 XX
 DT 15-JUN-1993 (first entry)
 DE Human tau-protein.
 KW Alzheimer's disease; diagnosis; subtyping; monitoring; assay.
 XX Homo sapiens.
 PN W09303369-A.
 PD 18-FEB-1993.
 PF 03-AUG-1992; 92WO-0506382.
 PR 01-AUG-1991; 91US-073878.
 XX (VOOR/) VOORHEIS P H.
 PI Voorheis PH;
 DR WPI; 1993-076670/09.
 DR N-PSDB; AAQ37305.
 XX
 PT Method for diagnosing, subtyping and monitoring Alzheimer's
 PT disease - by assaying a sample of body fluid for the presence of a
 PT tau-peptide using an anti-tau antibody
 XX Disclosure; Flg 1; 43pp; English.
 PS
 CC The sequence is that one form of human tau protein (from Goedert
 CC et al., PNAS USA 85: 4051-4055) which was used for the prodn.
 CC of anti-tau peptide antibodies. These are used as part of a method
 CC for diagnosing, subtyping or monitoring Alzheimer's disease by
 CC assaying a sample of body fluid for the presence of a tau-peptide
 CC using an anti-tau antibody or the presence of an anti-tau-peptide
 CC autoantibody. The methods can be used for confirming a clinical
 CC diagnosis of Alzheimer's disease and in following the course of the
 CC disease and treatment.
 XX
 SQ Sequence 352 AA;

Query Match 39.7%; Score 1574; DB 14; Length 352;
 Best Local Similarity 46.2%; Pred. No. 3.2e-90;
 Matches 350; Conservative 0; Mismatches 2; Indels 406; Gaps 4;

QY 1 MAEPQEFVEMDHAGTYGLGDRKDGTYTHQDQEGDITDAGLKESPLQTPPTDGESEPG 60
 DB 1 MAEPQEFVEMDHAGTYGLGDRKDGTYTHQDQEGDITDAGLKESPLQTPPTDGESEPG 44
 QY 61 SETSDAKSTTAEDVTAPLVDEGAPGKQAAQPHTEIPETCTTAEAGIGDTPTPSLEDEAAG 120
 DB 45 -----AEEAGIGDTPTPSLEDEAAG 62

QY 121 HVTQPESEKGVVQEGFLREPGPPGLSHQLMGMPAPLLPEGPREATRQPSGTGPEDETEG 180
 DB 63 HVTQ-----
 QY 181 GRHAPPELLKHOLLGLHOEGPPLKAGGKERPGSKKEEVEDRDVDDESSPQSPSKASPA 240
 DB 67 -----
 QY 241 QDGRPPQTAAREATSI PGFPAEGAIPLVDFLSKVSTEIPASEPDGSPSVGRAKQDAPLE 300
 DB 67 -----
 QY 301 FTFHVEITPNVQKEQAHSEHLAGRAFFGAPGEGPEARGLGEDTKREADLPPESEKQPA 360
 DB 67 -----
 QY 361 AAPRGKPVSRVQPKARVMVSKSKDGTGDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
 DB 67 -----ARMVSKSKDGTGDDKKA-----
 QY 421 DPLIQSSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLKGDGKTKIATPRGAAPGQK 480
 DB 85 -----KGDGKTKIATPRGAAPGQK 105
 QY 481 GQANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRSTPSLTPPTREP 540
 DB 106 GQANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRSTPSLTPPTREP 165
 QY 541 KKVAVVTPPKSPSAKSRLQTA PVMPDLKNVSKSGTSTENLKHQPGGKGVQIINKKLD 600
 DB 166 KKVAVVTPPKSASAKSRLQTA PVMPDLKNVSKSGTSTENLKHQPGGK----- 216
 QY 601 LSNVQSKGSDNIKHVPGGSGVQIVYKPVLDLSKVTSCGSLGNIHHKPGGQVEVKSEK 660
 DB 217 -----VQIVYKPVLDLSKVTSCGSLGNIHHKPGGQVEVKSEK 254
 QY 661 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720
 DB 255 LDFKDRVQSKIGSLDNIHVPVGGGKNIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 314
 QY 721 SPRHLSNVSSGSDIMVDSQPLATLADEVSASLAKOGL 758
 DB 315 SPRHLSNVSSGSDIMVDSQPLATLADEVSASLAKOGL 352

RESULT 13
 ABB57300
 ID ABB57300 standard; Protein; 364 AA.
 AC ABB57300;
 DT 07-MAR-2002 (first entry)
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:841.
 XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
 XX Mus musculus.
 XX WO200188188-A2.
 PD 22-NOV-2001.
 PF 18-MAY-2001; 2001WO-JP04192.
 PR 18-MAY-2000; 2000JP-0145977.
 XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 PI WPI; 2002-034733/04.
 DR

DR N-PSDB; ABI99761.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring

PT expression levels of particular genes defined in the specification or

PT by determining the expression profile of a gene group comprising these

PT genes -

XX

PS Claim 2; Page 2085-2086; 2690pp; English.

XX

CC The present invention describes a method for examining ischaemic

CC conditions, comprising measuring the expression levels of particular

CC genes (I) in a test sample or determining the expression profile of a

CC gene group in the sample comprising genes selected from (I). The method

CC is useful for examining the ischaemic condition (e.g. compressive

CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring

CC expression levels of particular genes (ABI99202 to ABI99912, encoding

CC the protein sequences in ABB57020 to ABB57374) or by determining the

CC expression profile of a gene group comprising these genes. The

CC expression levels or expression profiles produced by these genes are

CC used as an indicator when screening for ischaemic condition-improving

CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914

CC represent PCR primers for a mouse ischaemic condition related sequence,

CC which are used in the exemplification of the present invention.

XX

SQ Sequence 364 AA;

Query Match 34.1%; Score 1350.5; DB 23; Length 364;

Best Local Similarity 41.1%; Pred. No. 2.8e-76;

Matches 312; Conservative 13; Mismatches 13; Indels 421; Gaps 7;

QY 1 MAERQPEFVEDHAGTYGLGDRKQGGYTHMQDQGGTDAGLKESPLQTPTEDESGEPG 60

DB 1 MADPRQFDTMEDHA-----GDYTLQDQEGDMHGLK----- 33

QY 61 SETSDAKSTPTAEDYAPLVDEGAPGKAAQAQPHTEIPBGTAAEEAGIGDTSPLEDAAG 120

DB 34 -----AEEAGIGDTSPLEDAAG 120

QY 121 HVTQEPESGKYVQEGFLKEPGPGLSHQNSGMFGAPLLPEGPREATRQPSGTGPEDTEG 180

DB 52 HVTQ----- 55

QY 181 GRHAPELLKHOLLGDLHQEGPPLKAGGKERPKSKEVEDROVDVDESSPDSPSKASPA 240

DB 56 ----- 55

QY 241 QDGRPPQTAAREATSIQCFPAEGAIPLVDFLSKVSFTEIPASEPDGPGSVGRAGQDAPLE 300

DB 56 -----ARVA----- 59

QY 301 FTFHVEITPNVQKEQAHSEHLGRAAFPGAPGEGPARGPSLGEDTKREADLPPESEKQPA 360

DB 60 ----- 59

QY 361 AAPRGKPSVRVPQLKARWVSKSDGTGSDKKAKTSTRSSAKTLKRNPCLSKPLTPGSS 420

DB 60 -----SKDRTGNDEKKA----- 71

QY 421 DPLIQSSPAVCPPEPPSPKPHSVSTRTSGSGAKMKLKCADGKT--KIATPRGAAPP 478

DB 72 -----KGADGKTGAKIATPRGAASPA 92

QY 479 QKGQANATRIAPKTPPPSGEPKSGDRSGYSSPGSPGTPGSRKSRTPSLPTPTPTR 538

DB 93 QKGTSNATRIAPKTPPPSGEPKSGERSGYSSPGSPGTPGSRKSRTPSLPTPTPTR 152

QY 539 EPKKVAVRPPKSPSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKGVQIINKK 598

DB 153 EPKKVAVRTPPKSPSAKSLQATAPVMPDLKNVRSKIGSTENLKHQPGGK----- 205

QY 599 LDLSNVQSKGSKDNIRKHPVGGGVQIYKPVLDLSKVTSCGSLGNTHHHPGGGVQEVKS 658

DB 206 -----VQIYKPVLDLSKVTSCGSLGNTHHHPGGGVQEVKS 241

QY 659 EKLDKDRVQSKTGLSDNITHVPGGNGKKTETHTKLTFRENAKAKTOHGAEIVYKSPVWSG 718

DB 242 EKLDKDRVQSKTGLSDNITHVPGGNGKKTETHTKLTFRENAKAKTOHGAEIVYKSPVWSG 301

QY 719 DTSFRLHNSVSSGSDIMVDSPOLATLADSVASLAKQG 757

DB 302 DTSFRLHNSVSSGSDIMVDSPOLATLADSVASLAKQG 340

RESULT 14

AAV20956

ID AAY20956 standard; Protein; 314 AA.

XX

AC AAY20956;

XX

DT 22-JUL-1999 (first entry)

XX

DE Human microtubule associated protein BigTau wild type fragment 1.

XX

KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;

KW frameshift mutation; age-related disease; neurodegenerative disorder;

KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

KW Huntington's disease; multiple sclerosis; alcoholic liver disease;

KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;

KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;

KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;

KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;

KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;

KW high mobility group protein-C; neuroendocrine specific protein A.

XX

OS Homo sapiens.

XX

PN WO9845322-A2.

XX

PD 15-OCT-1998.

XX

PF 02-APR-1998; 98WO-IB00705.

XX

PR 10-APR-1997; 97US-0043163.

XX

PA (UYUT-) RIJKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYRO-) UNIV ROTTERDAM ERASMUS.

XX

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX

DR WPI; 1998-609901/51.

DR N-PSDB; AAY75763.

XX

XX

PT Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type

PT RNA

XX

PS Disclosure; Figure 12; 258pp; English.

XX

CC This invention describes a novel method for the diagnosis of a disease

CC caused by, or associated with, an RNA molecule that has a frameshift

CC mutation. The method is used to diagnose age-related diseases, especially

CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's

CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,

CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II

CC and many others listed) or susceptibility to these disorders. The method

CC allows a definitive diagnosis of Alzheimer's disease in living patients,

CC at an early stage. It is based on the observation that disease may be

CC caused by mutations in RNA rather than DNA. The invention describes the

CC used of neuronal system RNA molecules, specifically proteins including

CC beta-amyloid precursor protein (beta-APP), the microtubule associated

CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule

CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,

CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic

CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMP-C) and neuroendocrine specific protein A.

XX Sequence 314 AA;

Query Match 33.5%; Score 1329; DB 19; Length 314;

Best Local Similarity 100.0%; Pred. No. 5.1e-75;

Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 EPESKVVQEGFLRPPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPDTEGGRHA 184

DB 21 EPESKVVQEGFLRPPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPDTEGGRHA 80

QY 185 PELLKHQLGLDHOEGPLKAGGKERPGSKKEEVEDRDVDESSPQSPSPKASPAQDGR 244

DB 81 PELLKHQLGLDHOEGPLKAGGKERPGSKKEEVEDRDVDESSPQSPSPKASPAQDGR 140

QY 245 PPQTAAREATSIIPGPAAGAIPLVDFLSKVSTETIPASEPDGSPVGRAGKQDAPLETFH 304

DB 141 PPQTAAREATSIIPGPAAGAIPLVDFLSKVSTETIPASEPDGSPVGRAGKQDAPLETFH 200

QY 305 VEITENVQKEQAHSHEHLGRAAFGAPGEGPEARGPSLGEDTKADLPPEPSEKQPAAPR 364

DB 201 VEITENVQKEQAHSHEHLGRAAFGAPGEGPEARGPSLGEDTKADLPPEPSEKQPAAPR 260

QY 365 GKPSRVLPQK 375

DB 261 GKPSRVLPQK 271

RESULT 15

AAW23329

ID AAW23329 standard; Protein: 1831 AA.

AC AAW23329;

XX 23-MAR-1998 (first entry)

XX Microtubule-associated protein 2 (MAP2).

XX microtubule-associated protein 2; MAP2; neuroblastoma; human;

XX isoform; diagnosis; antibody; PCR primer; probe.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 1825

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX Misc-difference 1830

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

This is the human microtubule-associated protein 2 (MAP2). A 250 kD isoform of this 270 kD MAP2 is indicative of neuroblastoma. This isoform NB-MAP2 arises from alternative splicing of the primary MAP2 transcripts. Diagnostic agents for detecting NB-MAP2 in cells can be provided so as to detect and monitor neuroblastoma. The diagnostic agents are pairs of primers for specific amplification of DNA corresponding to at least part of the MAP2 transcript, provided neither primer binds to a sequence downstream of nucleotide 4168 or upstream of nucleotide 4510 of the present 5.5 kb sequence, or the corresponding antisense sequence, an oligonucleotide that binds to NB-MAP2-specific mRNA or the corresponding DNA in which nucleotides 4168 and 4510 are next to one another, an antibody specific for an epitope generated by translation of the region comprising the splice junction of NB-MAP2 specific mRNA, and a cDNA probe obtained by nick-translation or random priming of DNA or RNA specific for NB-MAP and obtainable from tissue or cell probes. The primers are used to detect MAP2-specific dimorphisms in tissue and cells by polymerase chain reaction (PCR) analysis. The oligonucleotide and the probes are used to detect NB-MAP2 specific transcripts (particularly by binding to membrane-bound mRNA) and the antibodies are used to detect NB-MAP2 in tissue and cells. Specifically these tests are used to detect neuroblastoma (or its precursor cells) and to monitor the success of bone marrow purging. The probes can detect bone marrow and lung metastases in cases where histopathological diagnosis of neuroblastoma is difficult. The oligonucleotide is used therapeutically to inhibit NB-MAP2 expression, resulting in development of neuroblastoma to normal ganglioneurons.

XX Sequence 1831 AA;

Query Match 20.5%; Score 812.5; DB 18; Length 1831;

Best Local Similarity 29.2%; Pred. No. 6.8e-42;

Matches 284; Conservative 120; Mismatches 279; Indels 291; Gaps 39;

QY 5 ROEFVEMEDHAGTYGLDRKDOGGYTMHQDQEGDTAGLKESPLQPTPTDGESEEPGSETS 64

DB 922 KDEFSV-DKEASAHISGDK---SGLSKFQDEKKAN-----DRLOFVLEK-SEEHADSKE 971

QY 65 DAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTAAEEA--GIGDTFSLDEAAGHV 122

DB 972 HAKKTEAGD-EIETRGLGVTYEQALAK-DLSIPTDASSEKAEKGLSSVPEI----- 1021

QY 123 TQPESSKVVQEGF-----LREPGLSLHQLASGM-----PCAPLLPEGPREA 166

DB 1022 -AEVPESSKVEQGLDFAVOGLDVKISDFG-----OMASGLNIDRRRATEKLEATQDM 1074

QY 167 TROPSTGTGPDPT-----EGGRHAPELLKHLQLGDL--LHQEGPPLK---GAGGK 209

DB 1075 T--PSSKAPQADAFMGVESHMKECTKYSETEVKQKVPDLVHQEAVDKESYESSGE 1132

QY 210 ERPGSKEEVDKRDVDESSPQDS-----PP--SKASPAQDGRP----- 245

DB 1133 HESLTMESLKADEGKETSPESSLIQDEIAVKLSVETPCPPAVSEADLATDERADVQNEF 1192

QY 246 ---PQTAAREATSIIPGPAAGAIPLVDFLSKVSTETIPASEPDGSPVGRAGKQDAPLEFT 302

DB 1193 IOGPKEESKETPDISITPTSDVAEPLHETIVSE--PAEIQSEEE-----IAQGEYDKLLER 1247

QY 303 F-----HVEITENVQK-----EQAHSSEHLGRAAFPG 329

DB 1248 SDTLQITDLGVSGAREFEVETCPSEHKGVIESVVTIEDDFITVVQTTTDEGESGSHVRF 1307

QY 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPSRVQPKAKRMVS 380

DB 1308 AALEQPEVERRPSPHDEEVEFEVEAAEAQAEPKDGSPASPSPERPEVA--LSEYKTKETD 1366

QY 381 KSKDGTGSDS-----KKAKTSTRSS-AKTLKNRP 408

DB 1367 DYKDETTIDDSIMDASLWVDTDQDDRSINTQETLPIPKAEKAEKARRSSLEKHKRKP 1426

QY 409 C-----LSPKPLTPGSSDPL-----IQPSSPA--VCPPEP 436

DB 1427 FKTRGRISTPPEKVKAKPESTVSRDEVRRKKAVYKAEKAEKAEKAEKAEKAEKAEKAE 1486

Diagnostic agents for detecting neuroblastoma specific isoform of microtubule-associated protein - for diagnosis or monitoring of neuroblastoma, are primers, probes or antibodies, also new therapeutic antisense sequence

Disclosure; Fig 2; 34pp; English.

Qy 437 ---SSPKHVSVTSRTSSGA-----KEMKLKAGDGKTKIATPRGAAP-----PG 478
Db 1487 IKYTRPHTLSCVKRKTAAAGGESALAPSVFKQAKDKVSDGVTKSPKRSLSLPRPSSILPP 1346
Qy 479 QKG-----QANATRIPAKT-----PAPKTPSS-----GEPPKSG 509
Db 1547 RRGVSGDRDENSFLNSISSARRTTRSEFIRRAGKSGTSTPTPGSTAITTEGTPESYS 1606
Qy 510 DRGYSSPGSPGTPGSRSTPSLPTPP-----TREPKKVAVVRTPPKSPSSAKSRLOQAP 564
Db 1607 SR---TPGTGCTP-SYPTPTHTGTPKSAIILVPSEKKVAIIRTPKASPGLTQKQLRLIN 1661
Qy 565 VPMPLKNVSKIGSTENLKHQPGGKVQIINKKLDLSNVQSKGSKDNIKHVPGGGSVQ 624
Db 1662 QPLPOLKNVSKIGSTONIKYQPKGGQVIVTKKIDLSH----- 1700
Qy 625 IVYKPVDSLKVTSCGSLGNHHPGGGQVEVKSEKLDKDRVQSKIGSLDNITHVPPGG 684
Db 1701 -----VTSKCGSLKNIRHPPGGGRVKIESVKLDKFEKAQAKYGSLDNAHHVPPGG 1750
Qy 685 NKIETHKLTPRENAKAKTDHGAEIVYKSPVVSQDTSRHLNVSVSTGSDMVDSPOLAT 744
Db 1751 NVKIDSOKLNPRENAKARVDHGAEIITQSPGRSSVAPRRLSNVSSGSINLLESPOLAT 1810
Qy 745 LADEVASLAKQGL 758
Db 1811 LAEDVTAALAKQGL 1824

Search completed: January 28, 2003, 14:14:11
Job time : 55.2376 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 9.93827 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQFEVEMDHAGTYGL.....SPQLATLADEVSASLAKOGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3965	100.0	758	10	US-09-904-987-5
2	2116.5	53.4	441	9	US-10-076-708-2
3	2116.5	53.4	441	9	US-10-066-810-1
4	2077.5	52.4	441	9	US-10-107-181-5
5	1776.5	44.8	383	9	US-10-076-708-4
6	1590	40.1	352	9	US-10-076-708-6
7	1590	40.1	391	10	US-09-790-148-23
8	733	18.5	140	9	US-10-107-181-9
9	575	14.5	109	9	US-10-107-181-3
10	561	14.1	108	9	US-10-107-181-2
11	546.5	13.8	109	9	US-10-107-181-1
12	546	13.8	140	9	US-10-107-181-10
13	364	9.2	67	10	US-09-790-148-1
14	285.5	7.2	1806	10	US-09-919-497-56
15	282	7.1	1325	10	US-09-864-761-35612
16	274	6.9	941	12	US-10-124-557-14
17	274	6.9	1022	12	US-10-124-557-84
18	274	6.9	1038	12	US-10-124-557-74
19	274	6.9	1049	12	US-10-124-557-58

20	274	6.9	1140	12	US-10-124-557-104	Sequence 104, Appl
21	274	6.9	1270	12	US-10-124-557-44	Sequence 44, Appl
22	274	6.9	1311	12	US-10-124-557-42	Sequence 42, Appl
23	274	6.9	1313	12	US-10-124-557-142	Sequence 142, Appl
24	274	6.9	1314	12	US-10-124-557-50	Sequence 50, Appl
25	274	6.9	1320	12	US-10-124-557-46	Sequence 46, Appl
26	274	6.9	1320	12	US-10-124-557-60	Sequence 60, Appl
27	274	6.9	1354	12	US-10-124-557-48	Sequence 48, Appl
28	274	6.9	1361	12	US-10-124-557-40	Sequence 40, Appl
29	274	6.9	1363	12	US-10-124-557-52	Sequence 52, Appl
30	274	6.9	1404	12	US-10-124-557-2	Sequence 2, Appl
31	274	6.9	1404	12	US-10-124-557-62	Sequence 62, Appl
32	247	6.2	1367	10	US-09-801-368-108	Sequence 108, Appl
33	243	6.1	113	10	US-09-864-761-44575	Sequence 44575, A
34	240	6.1	1848	9	US-09-839-936-6	Sequence 6, Appl
35	233	5.9	638	9	US-10-001-887-108	Sequence 108, Appl
36	229	5.8	1907	9	US-09-832-292-39	Sequence 39, Appl
37	228	5.8	2665	10	US-09-864-761-34248	Sequence 34248, A
38	227	5.7	826	10	US-09-894-998-47	Sequence 47, Appl
39	225.5	5.7	617	10	US-09-864-761-36182	Sequence 36182, A
40	223.5	5.6	1162	10	US-09-745-008-34	Sequence 34, Appl
41	222.5	5.6	881	10	US-09-816-860A-2	Sequence 2, Appl
42	221	5.6	755	10	US-09-919-497-57	Sequence 57, Appl
43	218	5.5	503	9	US-10-078-547-2	Sequence 2, Appl
44	218	5.5	802	10	US-09-823-240-2	Sequence 2, Appl
45	216.5	5.5	507	9	US-10-078-547-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-904-987-5

; Sequence 5, Application US/09904987

; Patent No. US20020037908A1

; GENERAL INFORMATION:

; APPLICANT: No. US20020037908A1actyl, Inc.

; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prep

; FILE REFERENCE: 42108/26146

; CURRENT APPLICATION NUMBER: US/09/904,987

; CURRENT FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 758

; TYPE: PRT

; ORGANISM: homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835

; DATABASE ENTRY DATE: 2001-02-13

; RELEVANT RESIDUES: (1)..(758)

; US-09-904-987-5

Query Match	100.0%	Score 3965;	DB 10;	Length 758;
Best Local Similarity	100.0%	Pred. No. 5.6e-180;		
Matches 758;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAEPQFEVEMDHAGTYGLGDRKQGGYTHQDQEGDTAGLKESPLQTPTEGSEPPG	60	
Db	1	MAEPQFEVEMDHAGTYGLGDRKQGGYTHQDQEGDTAGLKESPLQTPTEGSEPPG	60	
QY	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAQPHTEIEGTTAEAGTGDTPSLEDEAAG	120	
Db	61	SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAQPHTEIEGTTAEAGTGDTPSLEDEAAG	120	
QY	121	HYTQPESGKVVOEGFLREPGLSHOLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG	180	
Db	121	HYTQPESGKVVOEGFLREPGLSHOLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG	180	
QY	181	GRHAPELLKHQLLDLHQEGPLPKGAGGKERPGSKKEEVEDRDVDESSPQSPFSKASPA	240	
Db	181	GRHAPELLKHQLLDLHQEGPLPKGAGGKERPGSKKEEVEDRDVDESSPQSPFSKASPA	240	

Qy	241	QDGRPPOTAARENTSIPGPPAEGAIPLPVDFTLSKVSTEIPASSEPDPGSPVGRAKGDQADPLE	300
Db	241		
Qy	241	QDGRPPOTAARENTSIPGPPAEGAIPLPVDFTLSKVSTEIPASSEPDPGSPVGRAKGDQADPLE	300
Db	241		
Qy	301	FTFHVETIPNVQREQAHSSEHLGRAPFPAGPAGGPPARGPSLGEDTKADLPPESEKQPA	360
Db	301		
Qy	301	FTFHVETIPNVQREQAHSSEHLGRAPFPAGPAGGPPARGPSLGEDTKADLPPESEKQPA	360
Db	301		
Qy	361	AAPRGKPVSRVPOLKARMYSKSDGTGSDDKAKKTSTRSSAKTTLKNRPCLSPKLPDPGSS	420
Db	361		
Qy	361	AAPRGKPVSRVPOLKARMYSKSDGTGSDDKAKKTSTRSSAKTTLKNRPCLSPKLPDPGSS	420
Db	361		
Qy	421	DPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK	480
Db	421		
Qy	421	DPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK	480
Db	421		
Qy	481	GOANATRIPAKTTPAKTTPSSGEPKSGDRSGSYSPGSPGTPGSRSTRPSLPTPTREP	540
Db	481		
Qy	481	GOANATRIPAKTTPAKTTPSSGEPKSGDRSGSYSPGSPGTPGSRSTRPSLPTPTREP	540
Db	481		
Qy	541	KKVAVVTRTPKSPSSAKSRLQOTAPVMPDLKNVYSKIGSTENLKHQPGGKQVQIINKKLD	600
Db	541		
Qy	541	KKVAVVTRTPKSPSSAKSRLQOTAPVMPDLKNVYSKIGSTENLKHQPGGKQVQIINKKLD	600
Db	541		
Qy	601	LSNVQSKCGSKDNIKHVPGGGSVQIYVYKPVDSLKVTSKCGSLGNIHHKPEGGOVEVYKSEK	660
Db	601		
Qy	601	LSNVQSKCGSKDNIKHVPGGGSVQIYVYKPVDSLKVTSKCGSLGNIHHKPEGGOVEVYKSEK	660
Db	601		
Qy	661	LDPKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAKAKTDHGAELVYKSPVVSGDT	720
Db	661		
Qy	661	LDPKDRVQSKIGSLDNIHVPVGGNKKIETHKLTFFRENAKAKTDHGAELVYKSPVVSGDT	720
Db	661		
Qy	721	SPRHLNSVSTGSDMDVDSPLQATLADEYSASLAKOGL	758
Db	721		
Qy	721	SPRHLNSVSTGSDMDVDSPLQATLADEYSASLAKOGL	758
Db	721		
RESULT 2			
US-10-076-708-2			
; Sequence 2, Application US/10076708			
; Patent No. US20020164657A1			
; GENERAL INFORMATION:			
; APPLICANT: Sharma, Satish			
; APPLICANT: Rank, Kenneth			
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation			
; FILE REFERENCE: 6322			
; CURRENT APPLICATION NUMBER: US/10/076,708			
; CURRENT FILING DATE: 2002-02-15			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: Patentin version 3.0			
; SEQ ID NO 2			
; LENGTH: 441			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-076-708-2			
Query Match 53.4%; Score 2116.5; DB 9; Length 441;			
Best Local Similarity 58.2%; Pred. No. 3.1e-93;			
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps			
Qy	1	MAEPQREFEYWEHDAGTYGLGDRKDDGGTYTMHQDQEGDTDAGLKESPLQPTPTDGESEEPG	60
Db	1		
Qy	1	MAEPQREFEYWEHDAGTYGLGDRKDDGGTYTMHQDQEGDTDAGLKESPLQPTPTDGESEEPG	60
Db	1		
Qy	61	SETSDAKSTPTAEDVTAPLVDEGAPCKQAAAQPHTEIPSGTTAAEAGIGDTPSLDEAAG	120
Db	61		
Qy	61	SETSDAKSTPTAEDVTAPLVDEGAPCKQAAAQPHTEIPSGTTAAEAGIGDTPSLDEAAG	120
Db	61		
Qy	121*	HVTQEPESGKVVOEGFLREPGPPGLSHQLMSGMGPALLPEGPREATRQPSGTGPDTEG	180
Db	121		
Qy	121	HVTQ-----	124
Db	121	HVTQ-----	124

QY	181	GRHAPPELLKHQLLDLHQEGPPLKAGGKPERGSGKEEVDEDRVDDESSQDSPPSKASPA	241
DB	125	-----	124
QY	241	QDGRPPQTAAREATSIPGFAEGAIPLPVDLFSKVSTEIPASEPDGPSVGRAKGDAPLE	300
DB	125	-----	124
QY	301	FTFHVETTPNVQKEAHSEHLGRAAPGAPGEGPEARGLGDETKADLPPESEKQPA	360
DB	125	-----	124
QY	361	AAPRGKPVSRVPLQKARMVSKSDGTGSDDKAKATSTRSSAKTLKNRPCLSPKLPTPGSS	420
DB	125	-----	142
QY	421	DPLIQSSPAVCPPEPPSSPKHVSSTVSTRTOSSGAKEMKLGADGKTKIATPRGAAPPQOK	480
DB	143	-----	163
QY	481	GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSRPSSLPTPTPTREP	540
DB	164	GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSRPSSLPTPTPTREP	223
QY	541	KKVAVRTPPKSPSSAKSRLOTAPVMPMDLKNYKSKIGSTENLKHQPGGKVQIINKLD	600
DB	224	KKVAVRTPPKSPSSAKSRLOTAPVMPMDLKNYKSKIGSTENLKHQPGGKVQIINKLD	283
QY	601	LSNVQSCGSKDNIKHVPGGSGVQIYVKPVDLSKVTSKCGSLGNIHHKPGGGOVEVKSEK	660
DB	284	LSNVQSCGSKDNIKHVPGGSGVQIYVKPVDLSKVTSKCGSLGNIHHKPGGGOVEVKSEK	343
QY	661	LDFKDRVQSKIGSLDNTTHVPGGNGKKIETHKLTFFRENAKAKTDHGAEIVYKSPVVSQDT	720
DB	344	LDFKDRVQSKIGSLDNTTHVPGGNGKKIETHKLTFFRENAKAKTDHGAEIVYKSPVVSQDT	403
QY	721	SPRHLNSVSTGSDMDVDSPLQATLADEVASLAKOGL	758
DB	404	SPRHLNSVSTGSDMDVDSPLQATLADEVASLAKOGL	441
RESULT 3			
US-10-066-810-1			
; Sequence 1, Application US/10066810			
; Publication No. US20020188106A1			
; GENERAL INFORMATION:			
; APPLICANT: Mandelkow, Eva-Marla			
; Mandelkow, Eckhard			
; Lichtenberg-Kraag, Birgit			
; Biernat, Jacek			
; Drewes, Gerard			
; Steiner, Barbara			
; TITLE OF INVENTION: No. US20020188106A1e1 Tools For The Diagnostics			
; Treatment Of Alzheimer's Disease			
; NUMBER OF SEQUENCES: 1			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun			
; STREET: 233 South Waker Drive, 6300 Sears Tower			
; CITY: Chicago			
; STATE: Illinois			
; COUNTRY: United States of America			
; ZIP: 60606-6402			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Tape			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/066,810			
; FILING DATE: 04-Feb-2002			
; CLASSIFICATION: <Unknown>			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Joseph A. Williams, Jr.			

REGISTRATION NUMBER: 38,659
 REFERENCE/DOCKET NUMBER: 28384/32777
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-484-0448
 TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 441 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-066-810-1

Query Match 53.4%; Score 2116.5; DB 9; Length 441;
 Best Local Similarity 58.2%; Pred. No. 3.1e-93;
 Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPROFEFVEMDHAGTYGLDRKDGQGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
 Db 1 MAEPROFEFVEMDHAGTYGLDRKDGQGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120
 Db 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120

QY 121 HVTQEPESKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
 Db 121 HVTQ----- 124

QY 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDSESPQSPSKASPA 240
 Db 125 ----- 124

QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAGKQDAPLE 300
 Db 125 ----- 124

QY 301 FTFHVEITPNVQEQAHSEHLGRAFPAGCGEPEARGPSLGEDTKADLPPEPSEKQPA 360
 Db 125 ----- 124

QY 361 AAPRGKPVSRVQLKARVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTPGSS 420
 Db 125 ----- 142

QY 421 DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
 Db 143 ----- 163

QY 481 GOANATRIAPAKTPPAPKTPSSGEPKSGDRSGYSSPGTTPGSRSTRTPSLPTPTREP 540
 Db 164 GOANATRIAPAKTPPAPKTPSSGEPKSGDRSGYSSPGTTPGSRSTRTPSLPTPTREP 223

QY 541 KKVAVRTTPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGSGKVQIINKKLD 600
 Db 224 KKVAVRTTPPKSPSAKSLQATAPVMPDLKNVSKIGSTENLKHQPGSGKVQIINKKLD 283

QY 601 LSNVQSKGSKDNIIKHVPGGGSVQIVYKPVDSLKVTSCGSLGNIHHKPPGGQGVSEK 660
 Db 284 LSNVQSKGSKDNIIKHVPGGGSVQIVYKPVDSLKVTSCGSLGNIHHKPPGGQGVSEK 343

QY 661 LDFKDRVQSKIGSLDNIHVPGGKNKKTETHKLTFRENAKAKTDHGAIVYKSPVWSDT 720
 Db 344 LDFKDRVQSKIGSLDNIHVPGGKNKKTETHKLTFRENAKAKTDHGAIVYKSPVWSDT 403

QY 721 SPRHLNSVSTGSDIDWDSPLATLADSVASLAKOGL 758
 Db 404 SPRHLNSVSTGSDIDWDSPLATLADSVASLAKOGL 441

RESULT 4

US-10-107-181-5
 ; Sequence 5, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; STREET: Grenzachstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10107.181
 ; FILING DATE: 28-Mar-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913.915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 441 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 US-10-107-181-5

Query Match 52.4%; Score 2077.5; DB 9; Length 441;
 Best Local Similarity 57.5%; Pred. No. 2.1e-91;
 Matches 436; Conservative 0; Mismatches 5; Indels 317; Gaps 2;

QY 1 MAEPROFEFVEMDHAGTYGLDRKDGQGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60
 Db 1 MAEPROFEFVEMDHAGTYGLDRKDGQGYTMHQDQGGDTDAGLKESPLQTPPTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120
 Db 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPGTTAEAGIGDTPSLDEAAG 120

QY 121 HVTQEPESKVVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
 Db 121 HVTQ----- 124

QY 181 GRHAPELLKHQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDSESPQSPSKASPA 240
 Db 125 ----- 124

QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAGKQDAPLE 300
 Db 125 ----- 124

QY 301 FTFHVEITPNVQEQAHSEHLGRAFPAGCGEPEARGPSLGEDTKADLPPEPSEKQPA 360
 Db 125 ----- 124

QY 361 AAPRGKPVSRVQLKARVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTPGSS 420
 Db 125 ----- 142

QY 421 DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
 Db 143 ----- 163

QY 481 GOANATRIIPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRTPSLPTPTREP 540
Db 164 QGANATRIIPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRTPSLPTPTREP 223
QY 541 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVRTPPKSLSSAKSRLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKONIKVHPGGGSQVQIYKPVDSLKVTSKCGSLGNTHHKRPGGQVEVKSEK 660
Db 284 LSNVQSKGSKONIKQVPGGGSQVQIYKPVDSLKVTSKCGSLGNTHHKRPGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTTFRENAKAKTDHGAETVYKSPVVS GDT 720
Db 344 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTTFRENAKAKTDHGAETVYKSPVVS GDT 403
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVSASLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSPOLATLADEVSASLAKOGL 441

RESULT 5

US-10-076-708-4
; Sequence 4, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-4

Query Match 44.8%; Score 1776.5; DB 9; Length 383;
Best Local Similarity 50.5%; Pred. No. 2.7e-77;
Matches 383; Conservative 0; Mismatches 0; Indels 375; Gaps 3;

QY 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDQEGTDAGLKESPLQTPPTDGSSEPG 60
Db 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDQEGTDAGLKESPLQTPPTDGSSEPG 44
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
Db 45 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 62
QY 121 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTGPDTEG 180
Db 63 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTGPDTEG 66
QY 181 GRHAPELLKHQLLDLHQLGDLHQLGEPPLKGAGKERPGSKEEVEDRVDDESSPDSPSKASPA 240
Db 67 GRHAPELLKHQLLDLHQLGDLHQLGEPPLKGAGKERPGSKEEVEDRVDDESSPDSPSKASPA 66
QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
Db 67 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 66
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPAGAPGEPARGPSLGEDTKADLPPESEKQPA 360
Db 67 FTFHVEITPNVQEQAHSEHLGRAAFPAGAPGEPARGPSLGEDTKADLPPESEKQPA 66
QY 361 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
Db 67 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 84

QY 421 DPLIQSSPAVCEPPSPKHVSVTSGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 85 DPLIQSSPAVCEPPSPKHVSVTSGSSGAKEMKLGADGKTKIATPRGAAPPQK 105
QY 481 GOANATRIIPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRTPSLPTPTREP 540
Db 106 GOANATRIIPAKTPPSSGEPKSGDRSGYSSPGSGTSGRSRTPSLPTPTREP 165
QY 541 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 166 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 225
QY 601 LSNVQSKGSKDNKIHVPGGGSQVQIYKPVDSLKVTSKCGSLGNTHHKRPGGQVEVKSEK 660
Db 226 LSNVQSKGSKDNKIHVPGGGSQVQIYKPVDSLKVTSKCGSLGNTHHKRPGGQVEVKSEK 285
QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTTFRENAKAKTDHGAETVYKSPVVS GDT 720
Db 286 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTTFRENAKAKTDHGAETVYKSPVVS GDT 345
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVSASLAKOGL 758
Db 346 SPRHLSNVSTGSDIMVDSPOLATLADEVSASLAKOGL 383

RESULT 6

US-10-076-708-6
; Sequence 6, Application US/10076708
; Patent No. US20020164657A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Kenneth
; APPLICANT: Sharma, Satish
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-6

Query Match 40.1%; Score 1590; DB 9; Length 352;
Best Local Similarity 46.4%; Pred. No. 1.5e-68;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDQEGTDAGLKESPLQTPPTDGSSEPG 60
Db 1 MAEPQEEFVEMDHAGTYGLGDRKQGGYTMHQDQEGTDAGLKESPLQTPPTDGSSEPG 44
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
Db 45 SETSDAKSTPTAEDVTAPLVDEGAPKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 62
QY 121 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTGPDTEG 180
Db 63 HVTQEPESGKVQEGFLREPGLSHQLSMGMPGAPLLPEGPREATROPSTGTGPDTEG 66
QY 181 GRHAPELLKHQLLDLHQLGDLHQLGEPPLKGAGKERPGSKEEVEDRVDDESSPDSPSKASPA 240
Db 67 GRHAPELLKHQLLDLHQLGDLHQLGEPPLKGAGKERPGSKEEVEDRVDDESSPDSPSKASPA 66
QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
Db 67 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 66
QY 301 FTFHVEITPNVQEQAHSEHLGRAAFPAGAPGEPARGPSLGEDTKADLPPESEKQPA 360
Db 67 FTFHVEITPNVQEQAHSEHLGRAAFPAGAPGEPARGPSLGEDTKADLPPESEKQPA 66
QY 361 AAPRGKPVSRVQPKARVSKSKDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420

Db 67 -----ARMVSKSDGTGSDDKA----- 84
QY 421 DPLIQSSPAVCEPPSSPKHVSSTVSRGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
Db 85 -----KGADGKTKIATPRGAAPPGOK 105
QY 481 GOANATRIAPKTPPSSGPPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 540
Db 106 GOANATRIAPKTPPSSGPPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 165
QY 541 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 166 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 216
QY 601 LSNVQSKGSKDNKIHVPGGSGVQIYKVPDLKSVTSKCGSLGNTHHKKPGGQVEVKSEK 660
Db 217 -----VQIYKVPDLKSVTSKCGSLGNTHHKKPGGQVEVKSEK 254
QY 661 LDFKDRVQSKIGSLDNITHVPGGSKNKKIETHKLTFRNNAKAKTDHGAEIVYKSPVVSGDT 720
Db 255 LDFKDRVQSKIGSLDNITHVPGGSKNKKIETHKLTFRNNAKAKTDHGAEIVYKSPVVSGDT 314
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 315 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 352

RESULT 7

US-09-790-148-23
; Sequence 23, Application US/09790148
; Patent No. US20020001857A1
; GENERAL INFORMATION:
; APPLICANT: VANDERMEEREN, MARC;
; MERCKEN, MARC;
; VANMECHELEN, EUGEN;
; VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/790,148
; FILING DATE: 21-Feb-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: 08/256,167
; FILING DATE: 13-JUN-1994
; APPLICATION NUMBER: 08/244,951
; FILING DATE: 10-DEC-1993
; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 14-DEC-1992
; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410,003-1
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-790-148-23
Query Match 40.1%; Score 1590; DB 10; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.6e-68;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;
QY 1 MAEPQEFVEMDHAGTYGLGDRKDQGYTMHQDQEGDTAGLKESPLQPTEDGSEEPG 60
Db 40 MAEPQEFVEMDHAGTYGLGDRKDQGYTMHQDQEGDTAGLK----- 83
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAQPHTEPEGTTAEAGIGDTPSLEDAAG 120
Db 84 -----AEEAGIGDTPSLEDAAG 101
QY 121 HVTQEPESGKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
Db 102 HVTQ----- 105
QY 181 GRHAPELLKHOLLGDLHQEGPPLKGAGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
Db 106 ----- 105
QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGVSVGRAKGQDAPLE 300
Db 106 ----- 105
QY 301 FTFHVEITPNVQEQAHSEHGLRAAPGAPGEGPEARGPSLGEDTKEADLPSEKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVQDLKARMVSKSDGTGSDDKKAKTSTRSSAKTLKNRPLCLSPKLPFGSS 420
Db 106 -----ARMVSKSDGTGSDDKKA----- 123
QY 421 DPLIQSSPAVCEPPSSPKHVSSTVSRGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
Db 124 -----KGADGKTKIATPRGAAPPGOK 144
QY 481 GOANATRIAPKTPPSSGPPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 540
Db 145 GOANATRIAPKTPPSSGPPKSGDRSGYSSPGSGTTPGSRSTPSLPTPTREP 204
QY 541 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 600
Db 205 KKVAVRTPPKSSAKSRLQATAPVMPDLKNVSKIGSTENLKHQPGGKQVQIINKKLD 255
QY 601 LSNVQSKGSKDNKIHVPGGSGVQIYKVPDLKSVTSKCGSLGNTHHKKPGGQVEVKSEK 660
Db 256 -----VQIYKVPDLKSVTSKCGSLGNTHHKKPGGQVEVKSEK 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGSKNKKIETHKLTFRNNAKAKTDHGAEIVYKSPVVSGDT 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGSKNKKIETHKLTFRNNAKAKTDHGAEIVYKSPVVSGDT 353
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 354 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 391
RESULT 8
US-10-107-181-9
; Sequence 9, Application US/10107181
; Patent No. us20020168687A1
; GENERAL INFORMATION:


```

, REPLICATION NUMBER: 107/107/107
, FILING DATE: 28-Mar-2002
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US/08/913,915
, FILING DATE: <Unknown>
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 061-688 51 08
, TELEFAX: 061-688 13 95
, TELEX: 962292/965542 hlr ch
, INFORMATION FOR SEQ ID NO: 2:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 108 amino acids
, TYPE: amino acid
, STRANDEDNESS: <Unknown>
, TOPOLOGY: linear
, MOLECULE TYPE: peptide
, SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-107-181-2

Query Match 14.1%; Score 561; DB 9; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 108; Conservative 0; Mismatches 0; Indels 0

```


QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 60
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDN 676
 Db 61 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDN 108

RESULT 11

US-10-107-181-1
 ; Sequence 1, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; STREET: Grenzacherstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/107,181
 ; FILING DATE: 28-Mar-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 109 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-107-181-1
 Query Match 13.8%; Score 546.5; DB 9; Length 109;
 Best Local Similarity 77.9%; Pred. No. 5.5e-20;
 Matches 109; Conservative 0; Mismatches 0; Indels 31; Gaps 1;
 QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 29
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGNNKI 688
 Db 30 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGNNKI 89
 QY 689 ETHKLTFRNENAKAKTDHGAE 708
 Db 90 ETHKLTFRNENAKAKTDHGAE 109

RESULT 12

US-10-107-181-10
 ; Sequence 10, Application US/10107181
 ; Patent No. US20020168687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFFMANN-LA ROCHE AG
 ; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION
 ; NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
 ; STREET: Grenzacherstrasse 124
 ; CITY: Basle
 ; STATE: BS
 ; COUNTRY: Switzerland
 ; ZIP: CH-4070
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/107,181
 ; FILING DATE: 28-Mar-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/913,915
 ; FILING DATE: <Unknown>
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 061-688 51 08
 ; TELEFAX: 061-688 13 95
 ; TELEX: 962292/965542 hlr ch
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 140 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: <Unknown>
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-107-181-10
 Query Match 13.8%; Score 546; DB 9; Length 140;
 Best Local Similarity 71.4%; Pred. No. 7.6e-20;
 Matches 100; Conservative 23; Mismatches 17; Indels 0; Gaps 0;
 QY 569 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 628
 Db 1 DLKNNVSKIGSTENLKHQPGGKVQIINKKLDLSNNVQSKGSKDNKIKHVPGGSVQIYVK 60
 QY 629 PVDLSKVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGNNKI 688
 Db 61 KIDLSHVTSCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGNNKI 120
 QY 689 ETHKLTFRNENAKAKTDHGAE 708
 Db 121 DSQKLNFRNENAKAKTDHGAE 140

RESULT 13

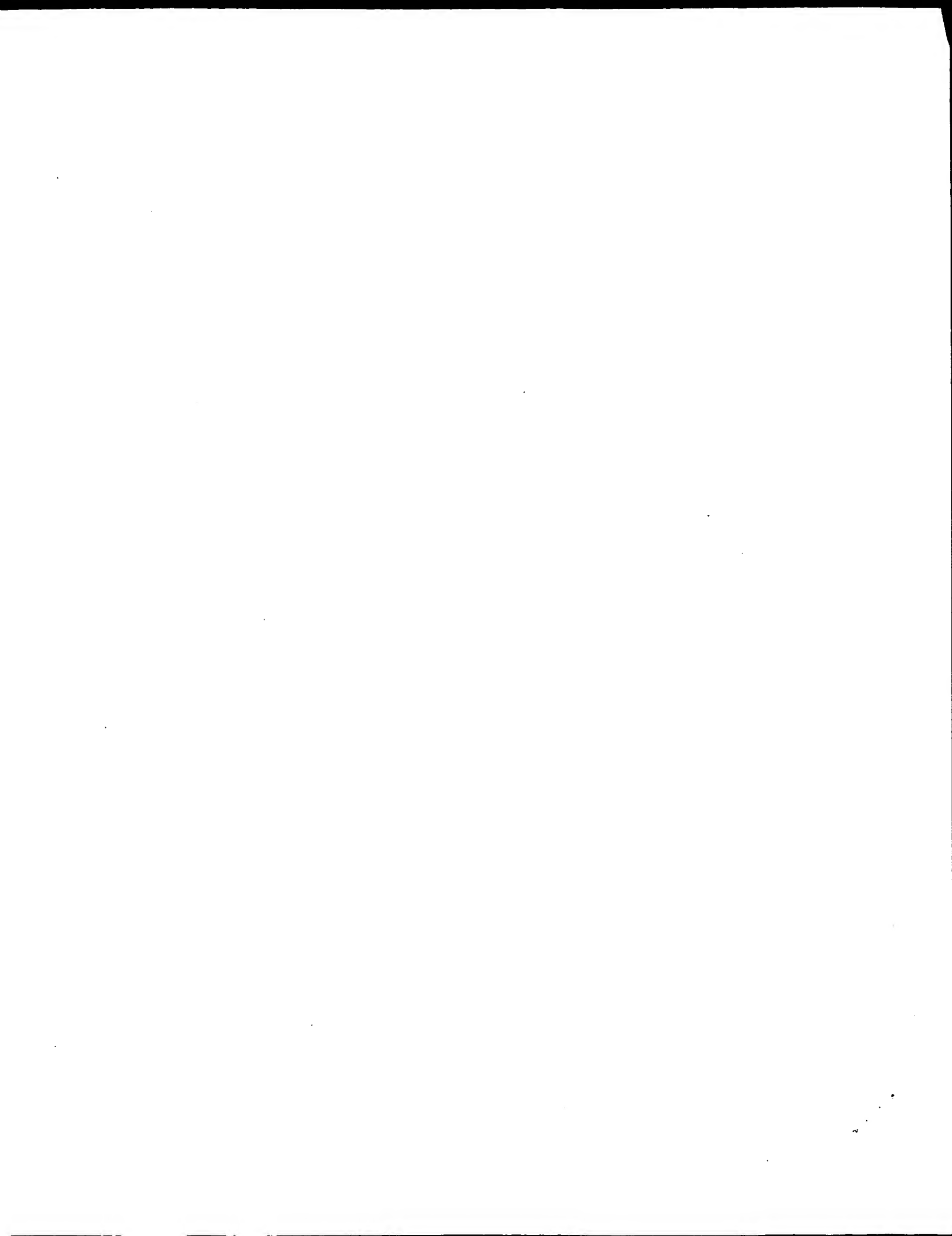
US-09-790-148-1
 ; Sequence 1, Application US/09790148
 ; Patent No. US20020001857A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VANDERMEEREN, MARC;
 ; MERCKEN, MARC;
 ; VANNEHELEN, EUGEN;
 ; VAN DE VOORDE, ANDRE
 ; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
 ; DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
 ; PROTEIN TAU, HYBRIDOMAS SECRETING THESE
 ; ANTIBODIES, ANTIGEN RECOGNITION BY THESE
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIERMAN & MUSERLIAN
 ; STREET: 600 THIRD AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10016
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/790,148
 ; FILING DATE: 21-Feb-2001
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: <Unknown>
 ; FILING DATE: 27-JUN-1994
 ; APPLICATION NUMBER: 08/256,167
 ; FILING DATE: 13-JUN-1994
 ; APPLICATION NUMBER: 08/244,951
 ; FILING DATE: 10-DEC-1993
 ; APPLICATION NUMBER: PCT/EP93/03499
 ; FILING DATE: 14-DEC-1992
 ; APPLICATION NUMBER: EP/92/403403.6
 ; FILING DATE: 14-DEC-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CHARLES A. MUSERLIAN
 ; REGISTRATION NUMBER: 19,683
 ; REFERENCE/DOCKET NUMBER: 410,003-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 661-8000
 ; TELEFAX: (212) 661-8002
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 67
 ; TYPE: Amino Acid
 ; STRANDEDNESS: Unknown
 ; TOPOLOGY: Unknown
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-790-148-1

Query Match 9.2%; Score 364; DB 10; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 472 RGAAPPQKQANATRIAPKTPPSPGPPKSGDRSGYSSPGSGTSGSRTRPS 531
 Db 1 RGAAPPQKQANATRIAPKTPPSPGPPKSGDRSGYSSPGSGTSGSRTRPS 60
 QY 532 LPTPPTR 538
 Db 61 LPTPPTR 67

RESULT 14
 US-09-919-497-56
 ; Sequence 56, Application US/09919497
 ; Patent No. US2002010662A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mutter, George L.
 ; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
 ; FILE REFERENCE: B0801/7225
 ; CURRENT APPLICATION NUMBER: US/09/919,497
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/221,735
 ; PRIOR FILING DATE: 2000-07-31
 ; NUMBER OF SEQ ID NOS: 100
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 56
 ; LENGTH: 1806
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (758)..(758)
 ; OTHER INFORMATION: Xaa = any amino acid
 ; NAME/KEY: UNSURE
 ; LOCATION: (809)..(809)
 ; OTHER INFORMATION: Xaa = any amino acid
 ; US-09-919-497-56
 Query Match 7.2%; Score 285.5; DB 10; Length 1806;
 Best Local Similarity 23.3%; Pred. No. 1.8e-06;

Matches 173; Conservative 53; Mismatches 270; Indels 245; Gaps 38;
 QY 16 GTYGLGDRKDGQGYTMHQDQEGDTDA---CLK--ESPLQTPTEGSEERGETS-----64
 Db 1012 GPQISLGRDGPAGLRGPFGRGLPCAQAGPLKGGEPGQGPVGS--PGEKGSACTAG 1069
 QY 65 --DAKSTPTAEDVTAPLVDEGAPKQAAAOPHTPEGTAAE-----AGIGDT 111
 Db 1070 PIGLRGRPGPOGPPGAGEKAGEK-----POGAGRQVQVGLPGPAGPAGS 1121
 QY 112 PSLDEAAGHVTTQEPESGKVVQVQEGFLREPPGLSHOLMSGMPGAPLLPGPRE-ATRPQ 170
 Db 1122 PG-EDGKGEIGEPGQKSGKGGKGGPPGPG--QCPVGAAGIAGDGPGRGQ 1175
 QY 171 SGT-GPDETEGGRHAPELLKHQLGLDLHQEGPP---LKGAGGKERPGSKKEEVEDRDVD 225
 Db 1176 QMFGQKGDGARGFP-----GPPGPIQLGL-----PGPGEKGENGDVG 1216
 QY 226 ESSQDQSPPSKASP-----ADGRPPPTAAREATSIPIGFPFAEAIPLPVDLFLSKVSTE 278
 Db 1217 PWGP-PGPPGPRGPGQNGADGPOGPPGSGVSGVGEKGEPEAGNPG-----1365
 QY 279 IPASEPDGPGSVGRAGQDAPLETFTHVEITPNVQKEAHSEHLGRAAFPCAPGEGPEAR 338
 Db 1266 -----PGEAGVGGPKGE-----RGEKGEAGPFGAAGP-PCAK 1296
 QY 339 GPSLGEDTKADLPPESEKQPAAPRGKP---VSRVPQLKARVSKSKDGTGSDDKKAK 394
 Db 1297 GPP-GDD-----GPKGNPGVPVGPFGDPGPPGELGPAGQGVGGD--KGE 1337
 QY 395 TSTRSSAKTLKNRCLSPKLPPTGSSDPLIQPSSPAVCPEPPS--SPKHVSSVTSRTGSS 452
 Db 1338 DG-----DPGQGPFG-----PSGEAGPPGPPGKRGPPGAAGABRGQEK 1377
 QY 453 GAK-EMKLKGDGDKTKATPRGAAPPQKQANATRIAPKT-----PPAKPT 498
 Db 1378 GAKGEAAGAEPPGTGTPVGPQG--PAGKPGPEGLRIGPVPVEQGLFGAAGQDGPQPMG 1435
 QY 499 PP-----SSGE-----PPKSGDRSGYSSPGSPCTPGSRSTRPSLTTPP 536
 Db 1436 PGLPLGLKGDPSKGEKHPGLIGLPGPEQGEKGDRLPGTQGSFGAKG-DGGIFGP- 1493
 QY 537 TREPKKVAVVRTPPKSPSSAKSRLQATAPVPMPLKNYKSTIGSTENLKHOPGGKGVQIIN 596
 Db 1494 -----AGPLGPPGPG-----LPGPQGPKNKGST-----GPAGKGD---1526
 QY 597 KKLDSNVQSKGSKDNKHKVHPGGSVQIVYKPYVDLSKVTSKCSGLNHHKPGGGQVEV 656
 Db 1527 -----SGLPFGPGP-----PGPPEVVIQPLPILSKKTKR-----HTE--GMQADA 1565
 QY 657 KSEKLDKDRVQSKIGSLDNI 677
 Db 1566 DDNILDYSDGMEIFGSLNSL 1586
 RESULT 15
 US-09-864-761-35612
 ; Sequence 35612, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aomicca-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:14:29 ; Search time 196.488 Seconds
(without alignments)
2487.216 Million cell updates/sec

Title: US-09-904-987-5
Perfect score: 3965
Sequence: 1 MAEPQFEVMEHDHAGTYGL.....SPQLATLADEVSASLAKGL 758

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main : *
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3965	100.0	758	1	PCT-US02-22032-5
2	3965	100.0	758	23	US-09-904-987-5
3	3965	100.0	758	27	US-60-406-385-228
4	2116.5	53.4	441	8	US-08-477-648-1
5	2116.5	53.4	441	8	US-08-484-674-1
6	2116.5	53.4	441	14	US-09-035-708-1

7	2116.5	53.4	441	14	US-09-035-708A-1	Sequence 1, Appli
8	2116.5	53.4	441	17	US-09-386-347-71	Sequence 71, Appli
9	2116.5	53.4	441	20	US-09-640-737-1	Sequence 1, Appli
10	2116.5	53.4	441	20	US-09-640-738-1	Sequence 1, Appli
11	2116.5	53.4	441	24	US-10-066-810-1	Sequence 2, Appli
12	2116.5	53.4	441	24	US-10-076-708-2	Sequence 2, Appli
13	2116.5	53.4	441	27	US-60-271-102-2	Sequence 2, Appli
14	2083.5	52.5	441	15	US-09-142-613-1	Sequence 1, Appli
15	2077.5	52.4	441	25	US-10-107-181-5	Sequence 5, Appli
16	1776.5	44.8	383	24	US-10-076-708-4	Sequence 4, Appli
17	1776.5	44.8	383	27	US-60-271-102-4	Sequence 4, Appli
18	1590	40.1	352	24	US-10-076-708-6	Sequence 6, Appli
19	1590	40.1	352	27	US-60-271-102-6	Sequence 6, Appli
20	1590	40.1	391	6	US-08-244-951-9	Sequence 9, Appli
21	1590	40.1	391	21	US-09-790-148-23	Sequence 23, Appli
22	1590	40.1	391	21	US-09-790-148A-23	Sequence 23, Appli
23	1586	40.0	351	27	US-60-087-557-7	Sequence 7, Appli
24	1552	39.1	374	26	US-10-087-557-7	Sequence 108, App
25	1253.5	31.6	365	21	US-09-791-537-7697	Sequence 7697, Ap
26	1119	28.2	239	21	US-09-760-469-1123	Sequence 1123, Ap
27	1119	28.2	239	26	US-10-216-583-1123	Sequence 1123, Ap
28	733	18.5	140	25	US-10-107-181-9	Sequence 9, Appli
29	621.5	15.7	1125	18	US-09-430-656-152	Sequence 152, App
30	621.5	15.7	1125	21	US-09-713-508-152	Sequence 152, App
31	621.5	15.7	1125	21	US-09-713-572-152	Sequence 152, App
32	621.5	15.7	1610	18	US-09-430-656-22	Sequence 22, Appli
33	621.5	15.7	1610	21	US-09-713-508-22	Sequence 22, Appli
34	621.5	15.7	1610	21	US-09-713-572-22	Sequence 1, Appli
35	593	15.0	112	18	US-09-432-285-1	Sequence 3, Appli
36	575	14.5	109	25	US-10-107-181-3	Sequence 2, Appli
37	561	14.1	108	25	US-10-107-181-2	Sequence 1, Appli
38	546.5	13.8	109	25	US-10-107-181-1	Sequence 10, Appli
39	546	13.8	140	25	US-10-107-181-10	Sequence 1124, Ap
40	495	12.5	322	21	US-09-760-469-1124	Sequence 1124, Ap
41	495	12.5	322	26	US-10-216-583-1124	Sequence 11805, A
42	471	11.9	700	20	US-09-614-150-11805	Sequence 11819, A
43	471	11.9	700	27	US-60-167-217-11819	Sequence 231, App
44	471	11.9	700	27	US-60-171-625-231	Sequence 9576, Ap
45	471	11.9	700	27	US-60-173-464-9576	

ALIGNMENTS

RESULT 1
PCT-US02-22032-5
; Sequence 5, Application PC/TUS0222032
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Protein Assembly or A
; FILE REFERENCE: 42108/34520
; CURRENT APPLICATION NUMBER: PCT/US02/22032
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 09/904,987
; PRIOR FILING DATE: 12 July 2001
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
PCT-US02-22032-5

Query Match 100.0%; Score 3965; DB 1; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAEPQFEVMEHDHAGTYGLGDRKDGQGYTHQDQEGTDAGLKESPLQPTTDCSEPG 60
|||||

Db 1 MAEPQREFEVMEDHAGTYGLDRKQGGYTMHQDQEGDTAGLKESTPLQTPTEDEGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Qy 181 GRHAPELLKHQLLDGLHOGEPPLKAGGKERPGSKKEVEDRDVDDESSPODSSPKASPA 240
Db 181 GRHAPELLKHQLLDGLHOGEPPLKAGGKERPGSKKEVEDRDVDDESSPODSSPKASPA 240
Qy 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Db 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Qy 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPPESEKQPA 360
Db 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPPESEKQPA 360
Qy 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
Db 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
Qy 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Qy 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTREP 540
Db 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTREP 540
Qy 541 KKVAVRTPPKSPSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQGGKQVQIINKLD 600
Db 541 KKVAVRTPPKSPSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQGGKQVQIINKLD 600
Qy 601 LSNVQSKCGKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Db 601 LSNVQSKCGKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Qy 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 720
Db 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 720
Qy 721 SPRHLNSVSTGSDMDVDSPOLATLADEVASLAKQGL 758
Db 721 SPRHLNSVSTGSDMDVDSPOLATLADEVASLAKQGL 758

RESULT 2

US-09-904-987-5
; Sequence 5, Application US/09904987
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepath
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904,987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 758
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / NM_016835
; DATABASE ENTRY DATE: 2001-02-13
; RELEVANT RESIDUES: (1)..(758)
US-09-904-987-5

Query Match 100.0%; Score 3965; DB 23; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAEPQREFEVMEDHAGTYGLDRKQGGYTMHQDQEGDTAGLKESTPLQTPTEDEGSEEPG 60
Db 1 MAEPQREFEVMEDHAGTYGLDRKQGGYTMHQDQEGDTAGLKESTPLQTPTEDEGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
Qy 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Db 121 HVTQEPESGVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATROPSTGTPEDTEG 180
Qy 181 GRHAPELLKHQLLDGLHOGEPPLKAGGKERPGSKKEVEDRDVDDESSPODSSPKASPA 240
Db 181 GRHAPELLKHQLLDGLHOGEPPLKAGGKERPGSKKEVEDRDVDDESSPODSSPKASPA 240
Qy 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Db 241 ODGRPPQTAAREATSIPIGFAEAGIPLPVDLFSKVSTEIPASEPDGSPVGRAGQDAPLE 300
Qy 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPPESEKQPA 360
Db 301 FTFHVEITPNVQKEQAHEEHLGRAAFPGAGPEARGPSLGEDTKEADLPPESEKQPA 360
Qy 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
Db 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
Qy 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Db 421 DPLIOPSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
Qy 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTREP 540
Db 481 GOANATRIAPKTPAPKTPSSGEPKSGDRSGYSSPGSPCTPGSRSTPSLPTPTREP 540
Qy 541 KKVAVRTPPKSPSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQGGKQVQIINKLD 600
Db 541 KKVAVRTPPKSPSAKSLQTPAPVMPDLKNVSKSTIGSTENLKHQGGKQVQIINKLD 600
Qy 601 LSNVQSKCGKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Db 601 LSNVQSKCGKDNKIHVPGGGSVQIVYKPVDSLKVTSKCGSLGNIHHKPGGGQVEVKSEK 660
Qy 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 720
Db 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 720
Qy 721 SPRHLNSVSTGSDMDVDSPOLATLADEVASLAKQGL 758
Db 721 SPRHLNSVSTGSDMDVDSPOLATLADEVASLAKQGL 758

RESULT 3

US-60-406-385-228
; Sequence 228, Application US/60406385
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS T
; TITLE OF INVENTION: INTERACT WITH SRC TYROSINE KINASES AND THE SRC TYROSINE KINAS
; TITLE OF INVENTION: CANCER CELLS
; FILE REFERENCE: D0273 PSP
; CURRENT APPLICATION NUMBER: US/60/406,385
; CURRENT FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 228
; LENGTH: 758
; TYPE: PRT

ORGANISM: Homo sapiens
US-60-406-385-228

```

Query Match      100.0%; Score 3965; DB 27; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.6e-210;
Matches 758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEPQEFVEVMDHAGTYGLGRKDGQGYTMHODQEGDTDAGLKESPLQTTEDGSEEPG 60
   |||||||
Db 1 MAEPQEFVEVMDHAGTYGLGRKDGQGYTMHODQEGDTDAGLKESPLQTTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
   |||||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120

QY 121 HVTQEPESGKVQVEGFLREPGLSHQLMGMPGAPLPPGEPREATRQPSGTGPDTEG 180
   |||||||
Db 121 HVTQEPESGKVQVEGFLREPGLSHQLMGMPGAPLPPGEPREATRQPSGTGPDTEG 180

QY 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
   |||||||
Db 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240

QY 241 QDGRPPQTAAREATSIIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
   |||||||
Db 241 QDGRPPQTAAREATSIIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300

QY 301 FTFHVETPNVQEQHSEHLGRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360
   |||||||
Db 301 FTFHVETPNVQEQHSEHLGRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360

QY 361 AAPRGKPVSRVLPOLKARVMSKSDGTGSDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
   |||||||
Db 361 AAPRGKPVSRVLPOLKARVMSKSDGTGSDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420

QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQOK 480
   |||||||
Db 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQOK 480

QY 481 GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSPGPGTSGRSRTPSLPTPTREP 540
   |||||||
Db 481 GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSPGPGTSGRSRTPSLPTPTREP 540

QY 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKVQIINKKLD 600
   |||||||
Db 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKVQIINKKLD 600

QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSKCGLGNTHHKPGGQGVKSEK 660
   |||||||
Db 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSKCGLGNTHHKPGGQGVKSEK 660

QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVSGDT 720
   |||||||
Db 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVSGDT 720

QY 721 SPRHLNVSTGSDIDWVDSPLATLADEVASLAKOGL 758
   |||||||
Db 721 SPRHLNVSTGSDIDWVDSPLATLADEVASLAKOGL 758

```

RESULT 4

```

US-08-477-648-1
: Sequence 1. Application US/08477648
: GENERAL INFORMATION:
: APPLICANT: Mandelkow, Eva-Maria
: APPLICANT: Mandelkow, Eckhard
: APPLICANT: Lichtenberg-Kraag, Birgit
: APPLICANT: Biernat, Jacek
: APPLICANT: Drewes, Gerard
: APPLICANT: Steiner, Barbara
: TITLE OF INVENTION: Novel Tools For The Diagnosis And
: TITLE OF INVENTION: Treatment Of Alzheimer's Disease
: NUMBER OF SEQUENCES: 1

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,648
FILING DATE: June 7, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Joseph A. Williams, Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 28384/32777
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-484-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-648-1

```

```

Query Match      53.4%; Score 2116.5; DB 8; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

```

```

QY 1 MAEPQEFVEVMDHAGTYGLGRKDGQGYTMHODQEGDTDAGLKESPLQTTEDGSEEPG 60
   |||||||
Db 1 MAEPQEFVEVMDHAGTYGLGRKDGQGYTMHODQEGDTDAGLKESPLQTTEDGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120
   |||||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPGTTAAEAGIGDTPSLEDEAAG 120

QY 121 HVTQEPESGKVQVEGFLREPGLSHQLMGMPGAPLPPGEPREATRQPSGTGPDTEG 180
   |||||
Db 121 HVTQ----- 124

QY 181 GRHAPELLKHOLLGDLHOEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
   ----- 124

QY 241 QDGRPPQTAAREATSIIPGPAEGAIPLVDFLSKVSTEIPASEPDGSPVGRAGQDAPLE 300
   ----- 124

QY 301 FTFHVETPNVQEQHSEHLGRAAFPGAPGEPARGPSLGEDTKADLPPESEKQPA 360
   ----- 124

QY 361 AAPRGKPVSRVLPOLKARVMSKSDGTGSDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
   |||||||
Db 361 AAPRGKPVSRVLPOLKARVMSKSDGTGSDDKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420

QY 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQOK 480
   |||||||
Db 421 DPLIQSPSPAVCPPEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTATPRGAAPPQOK 480

QY 481 GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSPGPGTSGRSRTPSLPTPTREP 540
   |||||||
Db 481 GOANATRIAPKTPAPKTPPSSGEPKSGDRSGYSPGPGTSGRSRTPSLPTPTREP 540

QY 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKVQIINKKLD 600
   |||||||
Db 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHOPGGGKVQIINKKLD 600

QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSKCGLGNTHHKPGGQGVKSEK 660
   |||||||
Db 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSKCGLGNTHHKPGGQGVKSEK 660

QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVSGDT 720
   |||||||
Db 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRENAKAKTDHGAETVYKSPVSGDT 720

QY 721 SPRHLNVSTGSDIDWVDSPLATLADEVASLAKOGL 758
   |||||||
Db 721 SPRHLNVSTGSDIDWVDSPLATLADEVASLAKOGL 758

```


QY 541 KKVAVVTRTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 224 KKVAVVTRTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 283
 QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNHKKHKKPGGQGVKSEK 660
 DB 284 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNHKKHKKPGGQGVKSEK 343
 QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTRENAKAKTDHGAIEIVYKSPVVS GDT 720
 DB 344 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTRENAKAKTDHGAIEIVYKSPVVS GDT 403
 QY 721 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
 DB 404 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 441

RESULT 5

US-08-484-674-1
 ; Sequence 1, Application US/08484674
 ; GENERAL INFORMATION:
 ; APPLICANT: Mandelkow, Eva-Maria
 ; APPLICANT: Mandelkow, Eckhard
 ; APPLICANT: Lichtenberg-Kraag, Birgit
 ; APPLICANT: Biernat, Jacek
 ; APPLICANT: Drewes, Gerard
 ; APPLICANT: Steiner, Barbara
 ; TITLE OF INVENTION: Novel Tools For The Diagnosis And
 ; TITLE OF INVENTION: Treatment Of Alzheimer's Disease
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Tape
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,674
 ; FILING DATE: June 7, 1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Joseph A. Williams, Jr.
 ; REGISTRATION NUMBER: 38,659
 ; REFERENCE/DOCKET NUMBER: 28384/32778
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-474-6300
 ; TELEFAX: 312-484-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 441 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-484-674-1

Query Match 53.4%; Score 2116.5; DB 8; Length 441;
 Best Local Similarity 58.2%; Pred. No. 3 2e-108;
 Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
 QY 1 MAEPQEFVEMDHAGTYGLGDRKDGQVTMHQDQEGDTAGLKESPLQTPTEDESGEPG 60
 DB 1 MAEPQEFVEMDHAGTYGLGDRKDGQVTMHQDQEGDTAGLKESPLQTPTEDESGEPG 60
 QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTEIPGCTTAEAGIGDTPSLEDAAG 120

DB 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHTEIPGCTTAEAGIGDTPSLEDAAG 120
 QY 121 HVTQEPESGKVVQEGFUREPQPGGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
 DB 121 HVTQ----- 124
 QY 181 GRHAPELLKHOLLGDLHQEGPPLKAGAGKERPGSKKEEVEDRDVDESSPODSSPKASPA 240
 DB 125 ----- 124
 QY 241 QDGRPPQTAAREATSIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGVGRAKGQDAPLE 300
 DB 125 ----- 124
 QY 301 FTFHVEITPNVQKQAHSEHLEHGRAAFPAGPEGEGEARGPSLGEDTKEADLPEPSEKQPA 360
 DB 125 ----- 124
 QY 361 AAPRGKPVSRYPQLKARVSKSKDGTGSDKKAKTSTRSSAKTLKNRPCLSPKLPPTGSS 420
 DB 125 -----ARVSKSKDGTGSDDKA----- 142
 QY 421 DPLIQSSPAVCEPPSPKHHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPCQK 480
 DB 143 -----KGADGKTKIATPRGAAPPCQK 163
 QY 481 QANATRIKAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 540
 DB 164 QANATRIKAPKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 223
 QY 541 KKVAVVTRTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 600
 DB 224 KKVAVVTRTPPKSPSSAKSLRLQAPVMPDLKNVSKIGSTENLKHQPGGKVQIINKKLD 283
 QY 601 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNHKKHKKPGGQGVKSEK 660
 DB 284 LSNVQSKGSKDNKIHVPGGGSGVQIYKPVDSLKVTSCGSLGNHKKHKKPGGQGVKSEK 343
 QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTRENAKAKTDHGAIEIVYKSPVVS GDT 720
 DB 344 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTRENAKAKTDHGAIEIVYKSPVVS GDT 403
 QY 721 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 758
 DB 404 SPRHLNSVSTGSDIMVDSPLATLADEVASLAKOGL 441

RESULT 6

US-09-035-708-1
 ; Sequence 1, Application US/09035708
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Cincinnati
 ; TITLE OF INVENTION: METHOD OF DETECTING AXONAL DAMAGE, ASSOCIATED DISEASE
 ; TITLE OF INVENTION: STATES, AND RELATED MONOCLONAL ANTIBODIES AND PROTEIN
 ; TITLE OF INVENTION: CONTROLS THEREFOR
 ; FILE REFERENCE: 1259-064
 ; CURRENT APPLICATION NUMBER: US/09/035,708
 ; CURRENT FILING DATE: 1999-03-05
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 441
 ; TYPE: PRT
 ; ORGANISM: human clone ht4u40 isoform reduced
 ; PUBLICATION INFORMATION:
 ; AUTHORS: Goedert, M.
 ; TITLE: Multiple Isoforms of Human Microtubule-Associated
 ; TITLE: Protein Tau: Sequences and Localization in
 ; TITLE: Neurofibrillary Tangles of Alzheimer's Disease
 ; JOURNAL: Neuron
 ; VOLUME: 3
 ; PAGES: 519-526
 ; DATE: OCT-1989

US-09-035-708-1

```
Query Match      53.4%; Score 2116.5; DB 14; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVEFVEDHAGTYGLGRKDGQGYTMHQDQGGDTAGLKESPLQTPTEGSEEPG 60
   |||||
Db 1 MAEPQEFVEFVEDHAGTYGLGRKDGQGYTMHQDQGGDTAGLKESPLQTPTEGSEEPG 60

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGTGTPTSLEDEAAG 120
   |||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGTGTPTSLEDEAAG 120

QY 121 HVTQEPESGKVQEGFLREPGLSHOLMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
   |||||
Db 121 HVTQ-----

QY 181 GRHAPELLKHOLLGDLHQEGPPLKGAGKERPGSKKEEVEDRDVDESSPDSPSSKASPA 240
   |||||
Db 125 -----

QY 241 QDGRPPQTAAREATSI PGPAEGAIPLVDFLSKVSTEIPASEPDGPGSVGRAKQODAPLE 300
   |||||
Db 125 -----

QY 301 FTFHVEITPNVQKQAHSEHLGRAAPPGAPGEGPEARGPSLGEDTKEADLPPESEKQPA 360
   |||||
Db 125 -----

QY 361 AAPRGKPVSRVQPKARWVSKSDGTGSDDKKAKTSTRSSAKTLKNRCLSPKLPTPGSS 420
   |||||
Db 125 -----ARMVSKSDGTGSDDKKA-----

QY 421 DPLIQSSPAVCEPPSPKHSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQK 480
   |||||
Db 143 -----KGADGKTKIATPRGAAPPQK 163

QY 481 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 540
   |||||
Db 164 QGANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSGTPGSRRTPSLPTPTREP 223

QY 541 KKVAVVTRTPKSPSSAKSRLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
   |||||
Db 224 KKVAVVTRTPKSPSSAKSRLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283

QY 601 LSNVQSKCGSKDNKIHVPGGGSVQIYVYPVDLSKVTSCGSLGNTHHKPGGQVEVKSEK 660
   |||||
Db 284 LSNVQSKCGSKDNKIHVPGGGSVQIYVYPVDLSKVTSCGSLGNTHHKPGGQVEVKSEK 343

QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVS GDT 720
   |||||
Db 344 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIVYKSPVVS GDT 403

QY 721 SPRHLSNVSSGTSDMVDSPQLATLADEVASLAKOGL 758
   |||||
Db 404 SPRHLSNVSSGTSDMVDSPQLATLADEVASLAKOGL 441

RESULT 7
US-09-035-708A-1
; Sequence 1, Application US/09035708A
; GENERAL INFORMATION:
; APPLICANT: University of Cincinnati
; TITLE OF INVENTION: METHOD OF DETECTING AXONAL DAMAGE, ASSOCIATED DISEASE
; TITLE OF INVENTION: STATES AND RELATED MONOCLONAL ANTIBODIES AND PROTEIN
; TITLE OF INVENTION: CONTROLS THEREFOR
; FILE REFERENCE: 1259-064
; CURRENT APPLICATION NUMBER: US/09/035.708A
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

; LENGTH: 441

; TYPE: PRT

; ORGANISM: human clone ht40 isoform reduced

; PUBLICATION INFORMATION:

; AUTHORS: Goedert, M.

; TITLE: Multiple Isoforms of Human Microtubule-Associated Protein Tau:

; TITLE: Sequences and Localization in Neurofibrillary Tangles of Alzheimer's

; TITLE: Disease

; JOURNAL: Neuron

; VOLUME: 3

; PAGES: 519-526

; DATE: OCT-1989

US-09-035-708A-1

Query Match

Best Local Similarity 58.2%; Pred. No. 3.2e-108;

Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;


```

; Sequence 71, Application US/09386347
; GENERAL INFORMATION:
; APPLICANT: Daniel M. Michaelson et al.
; TITLE OF INVENTION: PEPTIDES AND SUBSTANCES, METHODS AND
; TITLE OF INVENTION: DEVICES USING SAME FOR DIAGNOSING AND
; TITLE OF INVENTION: TREATING NEURODEGENERATIVE DISORDERS
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; OPERATING SYSTEM: Windows version 3.11
; SOFTWARE: Word for Windows version 2.0
; SOFTWARE: converted to an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/386,347
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 1710/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 572-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-386-347-71

```

```

Query Match 53.4%; Score 2116.5; DB 17; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

Qy 1 MAEPRQEFVEMDHAGTYGLGRKQGGYTMHQDQEGDTDAGLKESPLQPTTEDGSEEPG 60
Db 1 MAEPRQEFVEMDHAGTYGLGRKQGGYTMHQDQEGDTDAGLKESPLQPTTEDGSEEPG 60
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGQAAQPHTEIPEGTAAEAGIGDTPSLEDAAG 120
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGQAAQPHTEIPEGTAAEAGIGDTPSLEDAAG 120
Qy 121 HVTQEPESGKVQSGFLREPGLSLHQLMSGMPGAPLLPGLPREATRQPSGTGPEDETEG 180
Db 121 HVTQ----- 124
Qy 181 GRHAPELLKHOLLGLHQEGPPLKGAGKERPGSKEEVEDRDVDESSPDSPSKASPA 240
Db 125 ----- 124
Qy 241 QDGRPPQTAAREATSIQFPFAEGAIPLVDFLSKVSTEIPASEPDPGVSGRAKQDAPLE 300
Db 125 ----- 124
Qy 301 FTFHVEITPNVQKQAHSEHGLGRAFPAGPGEGPARGPSLGBDTKEADLPPESEKQPA 360
Db 125 ----- 124
Qy 361 AAPRGKPVSRVPQLKARMVSKSGDGTGDDKKAKTSTRSSAKTLKRNRPCLSPKLPPTGSS 420

```

```

Db 125 -----ARMVSKSGDGTGDDKKK----- 142
Qy 421 DPLIQSSPAVCEPPSPKHSVSTRTSGSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Db 143 -----KGADGKTKIATPRGAAPPGQK 163
Qy 481 QGANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP 540
Db 164 QGANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTREP 223
Qy 541 KKVAVVTPPKSPSSAKSRLQATAPVMPDLKNVSKSGTSTENLKHQPGGKQVLIINKKLD 600
Db 224 KKVAVVTPPKSPSSAKSRLQATAPVMPDLKNVSKSGTSTENLKHQPGGKQVLIINKKLD 283
Qy 601 LSNVQSKCGSKDNKHKVPGGSGVQIVYKVPDLKSVTKSGSLGNHGHKPGGQGVSEKSEK 660
Db 284 LSNVQSKCGSKDNKHKVPGGSGVQIVYKVPDLKSVTKSGSLGNHGHKPGGQGVSEKSEK 343
Qy 661 LDFKDRVQSKIGSLDNIHVPVGGGNKKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 720
Db 344 LDFKDRVQSKIGSLDNIHVPVGGGNKKIETHKLTFRENAKAKTDHGAIEIVYKSPVSGDT 403
Qy 721 SPRHLSNVSTGSDIMVDSPLATLADEVASLAKOGL 758
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVASLAKOGL 441

```

```

RESULT 9
US-09-640-737-1
; Sequence 1, Application US/09640737
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
; Lichtenberg-Kraag, Birgit
; Biernat, Jacek
; Drewes, Gerard
; Steiner, Barbara
; TITLE OF INVENTION: Novel Tools For The Diagnosis And
; Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/640,737
; FILING DATE: 17-Aug-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph A. Williams, Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 28384/32777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-484-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```


US-09-640-737-1

```
Query Match      53.4%; Score 2116.5; DB 20; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTEDGSEEPG 60
D 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTEDGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
D 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
QY 121 HVTQEPESGKVQVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
D 121 HVTQ----- 124
QY 181 GRHAPPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
D 125 ----- 124
QY 241 QDGRPPOTAAREATSIIPGPAEGAIPLVPDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
D 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHGLGRAAFPGEPEARGPSLGEDTKADLPSEKQPA 360
D 125 ----- 124
QY 361 AAPRGKPVSRVQPKARMVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
D 125 ----- 142
QY 421 DPLIQSPSPAVCEPEPPSPKHHVSSVTSRTSGSGAKEMKLKGADGKTKIATPRGAAPGQK 480
D 143 ----- 163
QY 481 GOANATRIAPKTPAPKTPPSSGPPPKSGDRSGYSSPGSGTSGSRSPSLPTPTREP 540
D 164 GOANATRIAPKTPAPKTPPSSGPPPKSGDRSGYSSPGSGTSGSRSPSLPTPTREP 223
QY 541 KKVAVRTPPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQGGKQVQIINKKLD 600
D 224 KKVAVRTPPKSPSSAKSLQTAAPVMPDLKNVKSIGSTENLKHQGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPGGSVQIYVKPVDLSKVTSKCSLGNIIHHKPGGQGVKSK 660
D 284 LSNVQSKGSKDNKIHVPGGSVQIYVKPVDLSKVTSKCSLGNIIHHKPGGQGVKSK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGNNKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 720
D 344 LDFKDRVQSKIGSLDNITHVPGGNNKKIETHKLTFRENAKAKTDHGAETVYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIDVSPQLATLADSVASLAKOGL 758
D 404 SPRHLSNVSTGSDIDVSPQLATLADSVASLAKOGL 441
```

RESULT 10

```
US-09-640-738-1
; Sequence 1, Application US/09640738
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
;             Mandelkow, Eckhard
;             Lichtenberg-Kraag, Birgit
;             Biernat, Jacek
;             Drewes, Gerard
;             Steiner, Barbara
; TITLE OF INVENTION: Novel Tools For The Diagnosis And
;                   Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/640,738
FILING DATE: 17-Aug-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Joseph A. Williams, Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 28384/32777
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-484-0448
TELEX: 25-3856
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-640-738-1
```

```
Query Match      53.4%; Score 2116.5; DB 20; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;

QY 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTEDGSEEPG 60
D 1 MAEPQEFVWEDHAGTYGLGDRKQGGYTHHQDEGDTDAGLKESPLQPTTEDGSEEPG 60
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
D 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDAAG 120
QY 121 HVTQEPESGKVQVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
D 121 HVTQ----- 124
QY 181 GRHAPPELLKHOLLGDLHQEGPPLKAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
D 125 ----- 124
QY 241 QDGRPPOTAAREATSIIPGPAEGAIPLVPDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE 300
D 125 ----- 124
QY 301 FTFHVEITPNVQKEAHSEHGLGRAAFPGEPEARGPSLGEDTKADLPSEKQPA 360
D 125 ----- 124
QY 361 AAPRGKPVSRVQPKARMVSKSDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
D 125 ----- 142
QY 421 DPLIQSPSPAVCEPEPPSPKHHVSSVTSRTSGSGAKEMKLKGADGKTKIATPRGAAPGQK 480
D 143 ----- 163
QY 481 GOANATRIAPKTPAPKTPPSSGPPPKSGDRSGYSSPGSGTSGSRSPSLPTPTREP 540
D 164 GOANATRIAPKTPAPKTPPSSGPPPKSGDRSGYSSPGSGTSGSRSPSLPTPTREP 223
```


QY 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
|||||
Db 224 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPVGGGSGVQIYKPVDSLKVTSSKGSIGNIHHKPGGQGVVEVYKSEK 560
|||||
Db 284 LSNVQSKGSKDNKIHVPVGGGSGVQIYKPVDSLKVTSSKGSIGNIHHKPGGQGVVEVYKSEK 343
QY 661 LDFKORVQSKIGSLDNTTHVPVGGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720
|||||
Db 344 LDFKORVQSKIGSLDNTTHVPVGGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 403
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
|||||
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 441
RESULT 11
US-10-066-810-1
; Sequence 1, Application US/10066810
; GENERAL INFORMATION:
; APPLICANT: Mandelkow, Eva-Maria
; Mandelkow, Eckhard
; Lichtenberg-Kraag, Birgit
; Biernat, Jacek
; Drewes, Gerard
; Steiner, Barbara
; TITLE OF INVENTION: Novel Tools For The Diagnosis And
; Treatment Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Waker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/066,810
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph A. Williams, Jr.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 28384/32777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-484-0448
; TELE: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-066-810-1

Query Match 53.4%; Score 2116.5; DB 24; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|
Db 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|
QY 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTABEAGIGDTPSLEDEAAG 120

Db 61 SETSDAKSTPTAEDVTAPLVDGAPGKQAAQPHTEIPEGTTABEAGIGDTPSLEDEAAG 120
|
QY 121 HVTQPESSKVVQEGFLRPGPGGLSHQMSGMPGAPLLPEGREATRQPSGTGPTDTEG 180
|
Db 121 HVTQ----- 124
|
QY 181 GRHAPELLKHOLLGDLHOEGPPLKGAGGKERPGSGKEEVEDRDVDESSPODPPSKASPA 240
|
Db 125 ----- 124
|
QY 241 QDGRPPQTAAREATSIPGFPAGGAIPLPVDFLSKVSTEIPASEPDGVSVGRAKQDAPLE 300
|
Db 125 ----- 124
|
QY 301 FTFHVEITPNVQKEQAHSEHLGRAAPGAPGEGPEARPSLGEDTKEADLPSEKQPA 360
|
Db 125 ----- 124
|
QY 361 AAPRGKPVSRVQLKARMSKSKDGTGDDKAKTSTRSSAKTLKNRCLSPKLPPTGSS 420
|
Db 125 -----ARMVSKSKDGTGDDKKA----- 142
|
QY 421 DPLIQSPSPAVCPPEPPSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK 480
|
Db 143 -----KGADGKTKIATPRGAAPGQK 163
|
QY 481 GOANATRIPAKTPPAKTPPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 540
|
Db 164 GOANATRIPAKTPPAKTPPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 223
|
QY 541 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
|
Db 224 KKVAVRTTPKSPSSAKSLQATAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
|
QY 601 LSNVQSKGSKDNKIHVPVGGGSGVQIYKPVDSLKVTSSKGSIGNIHHKPGGQGVVEVYKSEK 660
|
Db 284 LSNVQSKGSKDNKIHVPVGGGSGVQIYKPVDSLKVTSSKGSIGNIHHKPGGQGVVEVYKSEK 343
|
QY 661 LDFKORVQSKIGSLDNTTHVPVGGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 720
|
Db 344 LDFKORVQSKIGSLDNTTHVPVGGGSKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT 403
|
QY 721 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 758
|
Db 404 SPRHLSNVSTGSDIMVDSPLATLADEVSASLAKOGL 441
|
RESULT 12
US-10-076-708-2
; Sequence 2, Application US/10076708
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing Aa-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/10/076,708
; CURRENT FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-708-2

Query Match 53.4%; Score 2116.5; DB 24; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|
Db 1 MAEPQEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTDAGLKESPLQPTPTDGSSEPG 60
|

QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTTAEAGIGDTPTPSLEDAAG 120
|||||
Db 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQAQPHTEIPEGTTAEAGIGDTPTPSLEDAAG 120
|||||
QY 121 HVTQEPESGKVVQEGFLREPGLSHQLMGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
|||||
Db 121 HVTQ----- 124
QY 181 GRHAPELLKHQLLGLHQLGDLHQPGLKAGGKERPGSKEEVEDRDVDDESSPDSPSPKASPA 240
Db 125----- 124
QY 241 QDGRPQTAAAREATSIIPGPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKGDQAPLE 300
Db 125----- 124
QY 301 FTFHVEITPNVQKEQAISEEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPPESEKQPA 360
Db 125----- 124
QY 361 AAPRGKPVSRVLPQLKARVSKSDGTGDDKAKTSTRSSAKTLKRNPCLSPKLPTPGSS 420
Db 125----- 142
QY 421 DPLIQSSPAVCEPPSPKHSVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGQK 480
Db 143----- 163
QY 481 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 540
Db 164 GOANATRIPAKTPAPKTPSSGEPKSGDRSGYSSPGSPGTPGSRSTPSLPTPTPREP 223
QY 541 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 600
Db 224 KKVAVRTPPKSPSSAKSRLQTAAPVMPDLKNVKSIGSTENLKHQPGGKQVQIINKKLD 283
QY 601 LSNVQSKGSKDNKIHVPGGGVSQIVYKPVDSLKVTSKCSGLGNIHHPGGGQVEVKSEK 660
Db 284 LSNVQSKGSKDNKIHVPGGGVSQIVYKPVDSLKVTSKCSGLGNIHHPGGGQVEVKSEK 343
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKEIETHKLTFRENAKAKTDHGAIEIVYKSPVVS GDT 720
Db 344 LDFKDRVQSKIGSLDNITHVPGGNGKKEIETHKLTFRENAKAKTDHGAIEIVYKSPVVS GDT 403
QY 721 SPRHLNSVSTGSDMVDSPQLATLADEVSAKQGL 758
Db 404 SPRHLNSVSTGSDMVDSPQLATLADEVSAKQGL 441

RESULT 13

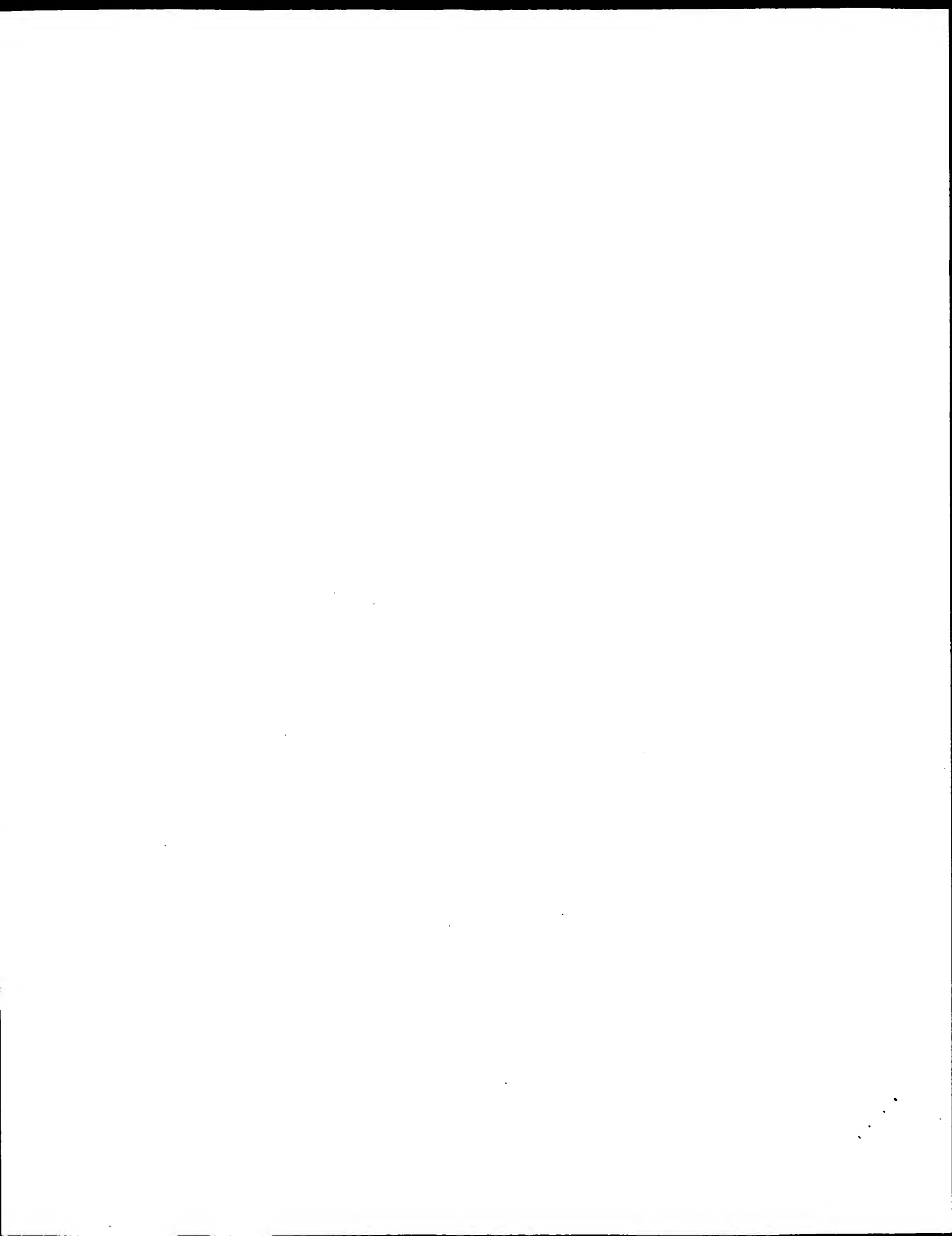
US-60-271-102-2
; Sequence 2, Application US/60271102
; GENERAL INFORMATION:
; APPLICANT: Sharma, Satish
; APPLICANT: Rank, Kenneth
; TITLE OF INVENTION: Assays for Accessing AA-Tau Aggregation
; FILE REFERENCE: 6322
; CURRENT APPLICATION NUMBER: US/60/271,102
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-271-102-2
Query Match 53.4%; Score 2116.5; DB 27; Length 441;
Best Local Similarity 58.2%; Pred. No. 3.2e-108;
Matches 441; Conservative 0; Mismatches 0; Indels 317; Gaps 2;
QY 1 MAEPQREFFVMEHAGTYGLGDRKQGGYTMHQDQGGDTAGLKESPLQPTPTDGSSEPG 60

RESULT 14

US-09-142-613-1
; Sequence 1, Application US/09142613A
; GENERAL INFORMATION:
; APPLICANT: ISHIGURO, Koichi
; APPLICANT: SATO, Kazuki
; APPLICANT: PARK, Jun-Mi
; APPLICANT: UCHIDA, Tsuneko
; APPLICANT: IMAHORI, Kazutomo
; TITLE OF INVENTION: ANTI-PHOSPHORYLATED TAU PROTEIN ANTIBODIES AND METHODS
; TITLE OF INVENTION: FOR DETECTING ALZHEIMER'S DISEASE WITH THE USE OF THE
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 98-0997*/LC(WMC)/1416
; CURRENT APPLICATION NUMBER: US/09/142,613A
; CURRENT FILING DATE: 1999-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens

Db 344 LDFKDRVQSKIGSLDNITHVPGGNKKIETHKLTVRENAKAKTDHGAELVYKSPVVS GDT 403
QY 721 SPRHLSNVSTGSDWVDSFQATLADSVASLAKQGL 758
Db 404 SPRHLSNVSTGSDWVDSFQATLADSVASLAKQGL 441

Search completed: January 28, 2003, 14:37:25
Job time : 202.488 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:16:13 : Search time 22.1541 seconds
(without alignments)
2701.343 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQFEVMEHDHAGTYGL.....SPQLATLADEVSLAKQGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 394613 seqs, 78952108 residues

Total number of hits satisfying chosen parameters: 394613

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1590	40.1	391	6	US-10-308-150-23
2	1400.5	35.3	340	5	US-09-724-676-49498
3	1400.5	35.3	340	5	US-09-724-676A-49498
4	1214	30.6	309	5	US-09-724-676-49497
5	1214	30.6	309	5	US-09-724-676A-49497
6	951.5	24.0	1856	5	US-09-724-676-51334
7	951.5	24.0	1856	5	US-09-724-676A-51334
8	891	22.5	502	5	US-09-724-676-51337
9	891	22.5	502	5	US-09-724-676A-51337
10	811	20.5	1825	5	US-09-724-676-51331
11	811	20.5	1825	5	US-09-724-676A-51331
12	750.5	18.9	471	5	US-09-724-676-51332
13	750.5	18.9	471	5	US-09-724-676A-51332
14	657	16.6	125	6	US-10-070-611-1
15	621.5	15.7	1125	6	US-10-100-957A-152
16	621.5	15.7	1125	6	US-10-100-957A-22
17	593	15.0	112	6	US-10-247-853-1
18	482.5	12.2	1011	1	PCT-US02-32851-39
19	470.5	11.9	94	6	US-10-070-611-10
20	462.5	11.7	94	6	US-10-070-611-13
21	456	11.5	91	6	US-10-070-611-12
22	455.5	11.5	94	6	US-10-070-611-11
23	415	10.5	1068	1	PCT-US02-32851-44
24	364	9.2	67	6	US-10-308-150-1
25	362.5	9.1	1693	5	US-09-724-676-51333
26	362.5	9.1	1693	5	US-09-724-676A-51333

ALIGNMENTS

RESULT 1

US-10-308-150-23

: Sequence 23, Application US/10308150

: GENERAL INFORMATION:

: APPLICANT: VANDEMEEREN, MARC; MERCKEN, MARC;

: VANMECHELEN, EUGEN;

: VAN DE VOORDE, ANDRE

: TITLE OF INVENTION: MONOCLONAL ANTIBODIES

: DIRECTED AGAINST THE MICROFUBULE-ASSOCIATED
: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
: ANTIBODIES, ANTIGEN RECOGNITION BY THESE

: NUMBER OF SEQUENCES: 24

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: BIERMAN & MUSERLIAN

: STREET: 600 THIRD AVENUE

: CITY: NEW YORK

: STATE: NEW YORK

: COUNTRY: USA

: ZIP: 10016

: COMPUTER READABLE FORM:

: MEDIUM TYPE: FLOPPY DISK

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: ASCII

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/10308,150

: FILING DATE: 02-Nov-2002

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: <Unknown>

: FILING DATE: 07-JUL-1999

: APPLICATION NUMBER: 09/348,952

: FILING DATE: 19-JAN-1995

: APPLICATION NUMBER: 08/403,917

: FILING DATE: 19-JAN-1995

: APPLICATION NUMBER: 08/403,916

: FILING DATE: 13-JUN-1994

: APPLICATION NUMBER: 08/244,951

: FILING DATE: 10-DEC-1993

: APPLICATION NUMBER: PCT/EP93/03499

: FILING DATE: 14-DEC-1992

: APPLICATION NUMBER: EP/92/403403.6

: FILING DATE: 14-DEC-1992

: ATTORNEY/AGENT INFORMATION:

: NAME: CHARLES A. MUSERLIAN

: REGISTRATION NUMBER: 19,683

: REFERENCE/DOCKET NUMBER: 410.003-1

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (212) 661-8000

: TELEFAX: (212) 661-8002

Sequence 51336, A
Sequence 51336, A
Sequence 1, Appli
Sequence 72859, A
Sequence 72859, A
Sequence 7237, Ap
Sequence 202, App
Sequence 202, App
Sequence 82068, A
Sequence 82068, A
Sequence 82069, A
Sequence 72, Appl
Sequence 405, App
Sequence 492, App
Sequence 12693, A
Sequence 75580, A
Sequence 75580, A
Sequence 139, App

27 302 7.6 339 5 US-09-724-676-51336
28 302 7.6 339 5 US-09-724-676A-51336
29 293 7.4 1374 1 PCT-US02-34846-1
30 288 7.3 1610 5 US-09-724-676-72859
31 288 7.3 1610 5 US-09-724-676A-72859
32 287.5 7.3 99 5 US-09-513-999C-7237
33 285.5 7.2 1806 1 PCT-US02-29560-202
34 285.5 7.2 1806 6 US-10-245-882-202
35 285 7.2 1054 5 US-09-724-676-82068
36 285 7.2 1054 5 US-09-724-676A-82068
37 285 7.2 1149 5 US-09-724-676-82069
38 285 7.2 1149 5 US-09-724-676A-82069
39 283.5 7.2 1466 6 US-10-257-021-72
40 283.5 7.2 1745 5 US-09-949-002-405
41 283.5 7.2 1771 5 US-09-949-002-492
42 282 7.1 1325 6 US-10-203-138A-12693
43 278.5 7.0 1636 5 US-09-724-676-75580
44 278.5 7.0 1636 5 US-09-724-676A-75580
45 277 7.0 1604 7 US-60-423-552-139

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-308-150-23

Query Match 40.1%; Score 1590; DB 6; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.7e-64;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;
QY 1 MAEPQEFVEMEDHACTYGLGRKDGQYTMHQDQGDGTDAGLKESPLQPTPTDGESEPG 60
Db 40 MAEPQEFVEMEDHACTYGLGRKDGQYTMHQDQGDGTDAGLKESPLQPTPTDGESEPG 83
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKAAQPHTEIPGTTAERAGIGDTPSLEDEAAG 120
Db 84 -----AEEAGIGDTPSLEDEAAG 101
QY 121 HVTQEPESGKVQEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPDTEG 180
Db 102 HVTQ----- 105
QY 181 GRHAPELLKQLLDLHQEGPLKAGGKERPGSKKEEVEDRDVDESSPDSPSKASPA 240
Db 106 ----- 105
QY 241 QDGRPPQTAAREATSTPGFPAEGAIFLPVDFLSKVSTEIPASEPDGPGSVGRAGQDAPLE 300
Db 106 ----- 105
QY 301 FTFHVEITPVQEQAHSEHLGRAAPGAPGEGPEARGPSLGEDTKADLPPESEKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVLPOLKARWVSKSDGTGSDDKAKKTSTRSAKTLKNRPCLSPKLPPTGSS 420
Db 106 -----ARMVSKSDGTGSDDKKA----- 123
QY 421 DPLIQSSPAVCPEPPSSPKHVSSVTSRTGSSGAKEMKLGADGKTKIATPRCAAPPQOK 480
Db 124 -----KGADGKTKIATPRCAAPPQOK 144
QY 481 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTPGSRSTPSPSLPTPTPREP 540
Db 145 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTPGSRSTPSPSLPTPTPREP 204
QY 541 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGKVOIINKLKD 600
Db 205 KKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENLKHQPGGK----- 255
QY 601 LSNVQSCGSKDNKHKVPGGGSQIVYKPVDSLKVTSKCSLGNIIHKKPGGQVEVKSEK 660
Db 256 -----VQIVYKPVDSLKVTSKCSLGNIIHKKPGGQVEVKSEK 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRFENAKAKTDHGAIEIVYKSPVVSGET 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRFENAKAKTDHGAIEIVYKSPVVSGET 353
QY 721 SPRHLNSVSTGSDIMVDSPLATLADSVASLAKOGL 758
Db 354 SPRHLNSVSTGSDIMVDSPLATLADSVASLAKOGL 391

RESULT 2
US-09-724-676-49498
; Sequence 49498, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49498
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-49498

Query Match 35.3%; Score 1400.5; DB 5; Length 340;
Best Local Similarity 75.1%; Pred. No. 4.8e-56;
Matches 280; Conservative 5; Mismatches 7; Indels 81; Gaps 3;
QY 353 EPSEKQPAAPRGKPVSRVPQL-----KARMVSKSDGTGSDDKAKKTSTRSAAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTCARMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPPTGSSDPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTKIATPRGAAPCGKGOANATRIAPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPCTP 523
Db 67 GKTKIATPRGAAPCGKGOANATRIAPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPCTP 126
QY 524 GSRSTPSLPTPTREPKKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPKKVAVRTPPKSPSSAKSRLOTAPVMPDLKNVSKIGSTENL 186
QY 584 KHQPGGKVOIINKLDSLNVQSKCGSKDNKHKVPGGGSQIVYKPVDSLKVTSKCSLGL 643
Db 187 KHQPGGKVOIINKLDSLNVQSKCGSKDNKHKVPGGGSQIVYKPVDSLKVTSKCSLGL 246
QY 644 NIHHKPGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRFENAKAKT 703
Db 247 NIHHKPGGQVEVKSEKLDKDRVQSKIGSLDNITHVPGGNGKKIETHKLTFRFENAKAKT 306
QY 704 DHGAEIVYKSPV 716
Db 307 DHGAEIVYKSPV 319

RESULT 3
US-09-724-676A-49498
; Sequence 49498, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49498
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-49498

Query Match 35.3%; Score 1400.5; DB 5; Length 340;
Best Local Similarity 75.1%; Pred. No. 4.8e-56;
Matches 280; Conservative 5; Mismatches 7; Indels 81; Gaps 3;
QY 353 EPSEKQPAAPRGKPVSRVPQL-----KARMVSKSDGTGSDDKAKKTSTRSAAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTCARMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLPPTGSSDPLIQSSPAVCPEPPSPKHVSSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTKIATPRGAAPCGKGOANATRIAPAKTPPAPKTPPSSGEPKSGDRSGYSSPGSPCTP 523

Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGGKQVQIINKLDLSNVQSKGSKDNIIKHVPGGGSVQIYKPVDSLKVTSCGSLG 643
Db 187 KHOPGGGKQVQIINKLDLSNVQSKGSKDNIIKHVPGGGSVQIYKPVDSLKVTSCGSLG 246
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 703
Db 247 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 306
QY 704 DHGAEIIVKSPV 716
Db 307 DHGAEIIVKSPV 319

RESULT 4

US-09-724-676-49497
; Sequence 49497, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49497
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-49497

Query Match 30.6%; Score 1214; DB 5; Length 309;
Best Local Similarity 66.8%; Pred. No. 1e-47;
Matches 249; Conservative 5; Mismatches 7; Indels 112; Gaps 4;

QY 353 EPSEKQPAAPRGKPVSRVQPL-----KARVMSKSDGTGSDDKKAKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTQARMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLTTPGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 523
Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGGKQVQIINKLDLSNVQSKGSKDNIIKHVPGGGSVQIYKPVDSLKVTSCGSLG 643
Db 187 KHOPGGGK-----VQIYKPVDSLKVTSCGSLG 215
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 703
Db 216 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 275
QY 704 DHGAEIIVKSPV 716
Db 276 DHGAEIIVKSPV 288

RESULT 5

US-09-724-676A-49497
; Sequence 49497, Application US/09724676A

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49497
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-49497

Query Match 30.6%; Score 1214; DB 5; Length 309;
Best Local Similarity 66.8%; Pred. No. 1e-47;
Matches 249; Conservative 5; Mismatches 7; Indels 112; Gaps 4;

QY 353 EPSEKQPAAPRGKPVSRVQPL-----KARVMSKSDGTGSDDKKAKTSTRSSAKT 403
Db 19 QPAEAEAG-----IGDTPSLEDEAAGHVTQARMVSKSDGTGSDDKKA----- 62
QY 404 LKNRPCLSPKLTTPGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGAD 463
Db 63 -----KGAD 66
QY 464 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 523
Db 67 GKTATPRGAAPPQGGKQANATRIAPKTPPSSSEPPKSGDRSGYSSPGSGTP 126
QY 524 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 583
Db 127 GSRSTPSLPTPTREPCKKAVVVRTPPKSPSSAKSLRQAPVMPDLKNVSKIGSTENL 186
QY 584 KHOPGGGKQVQIINKLDLSNVQSKGSKDNIIKHVPGGGSVQIYKPVDSLKVTSCGSLG 643
Db 187 KHOPGGGK-----VQIYKPVDSLKVTSCGSLG 215
QY 644 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 703
Db 216 NIHKPGGGQVEVSKELDFKDRVQSKIGSLDNITHVPGGKNKIIETHKLTFRENAKAKT 275
QY 704 DHGAEIIVKSPV 716
Db 276 DHGAEIIVKSPV 288

RESULT 6

US-09-724-676-51334
; Sequence 51334, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51334
; LENGTH: 1856
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51334

Query Match 24.0%; Score 951.5; DB 5; Length 1856;
Best Local Similarity 31.5%; Pred. No. 3.9e-35;
Matches 307; Conservative 125; Mismatches 281; Indels 261; Gaps 39;

QY 5 RQEFVEMDHAGTYGLGDRKQGGYTHHQDQEGDTDAGLKESPLQTPTEGSEBPGSETS 64
Db 924 KDEFV-DKEASAHISGDK---SGLSKEFDQEKKAN-----DRDVTLEK-SEHADSKE 973
QY 65 DAKSTPTAEDYATPLVDEGAPGKQAAQPHTEIPGTTAEEA--GIGDTPSLEDEAAGHV 122


```
Db 974 HAKTEAGD-EIETFGVLYEQALAK-DLSIPTDASSEKAEGKGLSSVPEI----- 1023
Qy 123 TQPESEKVVQEGF-----LREPGLSHQLMSGM-----PGAPLLPEGPREA 166
Db 1024 -AEVPSKVEOGLDFAVQGLDVKISDFG-----QMASGLNIDRRATELKLQATQDM 1076
Qy 167 TROPSTGTPEDT-----EGGRHAPPELLKHQLGDD--LHQEGPPLK---GAGGK 209
Db 1077 T--PSSKAPQADAFMGVESHMKEGTKVSETEVKQKAPDLVHQEAVDKVEESYSSGE 1134
Qy 210 ERPGSKEEVEDRDVDSESSPQDS-----PP--SKASPAQDGRP----- 245
Db 1135 HESLTMESLKADEGKETSPESLIQDEIAVKLSVEIPCPVAVSEADLATDRADVQMEF 1194
Qy 246 ---PQTAAREATSPGPAEAGIPLVPDLFSLKVSVEIPASEPDGPGSVGRAKGQDAPLEFT 302
Db 1195 IQPKESKETPDITSPDVAEPLHETIVSE-PAEQSEEEE-----IEAQGEYDKLLFR 1249
Qy 303 F-----HVEITPNVQK-----EQAHSEHGLGRAAFP 329
Db 1250 SDTLQITDLGVSAREFVETCPSEHKGVIESTIEDDFTVVQVTTTDEGESGSHSVRF 1309
Qy 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPVSRVPOLKARVMS 380
Db 1429 EKTGRGRISTPERKVAKEPSTVSRDEVRKRAVYKAEAKKTEVQAHSPSRKFLKPA 1488
Qy 437 ---SSPKHVSVTSRTGSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1489 IKYTRPHTLSCVKKRTTAAGGESALAPSVFKQAKDKVSDGVTSPKRSRLPRSSILPP 1548
Qy 479 OKG-----QANATRIPAKTP-----PAPKTPPSS---GEPPKSG 509
Db 1549 RRGVSGDRDENSEFSLNSSISSARRTTRSEPIRRAGKSGTSTPTTGTATPTGTPPSS 1608
Qy 510 DRGYSYSPGSGTSGSRTPSLTPPP-----TREPKVAVVTRTPPKSPSAKSLRQTAP 564
Db 1609 SR---TPGTGTP-SYRTPHTPGTKSAILVPSEKKVAIIRTPPKSPATPK-QLRLIN 1662
Qy 565 VPMPLKKNVSKIGSTENLKHQGGGVQIINKKLDLSNVQSKGSKDNIKHVPVGGG 624
Db 1663 QPLPDLKKNVSKIGSTDNIKYQPKGGQVRLINKKIDFSKVQRCGSKDNIKHAGGNNVQ 1722
Qy 625 IVYKVPDLKSVTSKCGSLGNIIHKPGQGVQVSEKLDKDRVQSKIGSLDNIHVPVGGG 684
Db 1723 IVTKKIDLSHVTSCGSLANIRHRRGGVRKTESVKLDFEKAQAQKVGSLDNAHVPGG 1782
Qy 685 NKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVGDTSPRHLNSVSTGSDIMVDSPLQAT 744
Db 1783 NVKIDSKLNFREHAKARVDHGAIEITQPGRSVASPRRLSNVSSGSLNLESPLQAT 1842
Qy 745 LADEVASLAKOGL 758
Db 1843 LAEDVTAALAKOGL 1856
```

RESULT 7

```
US-09-724-676A-51334
; Sequence 51334, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
```

```
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51334
; LENGTH: 1856
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51334
```

Query Match 24.0%; Score 951.5; DB 5; Length 1856;

Best Local Similarity 31.5%; Pred. No. 3.9e-35;
Matches 307; Conservative 125; Mismatches 281; Indels 261; Gaps 39;

```
Qy 5 ROBEVEVEDIAGTYGLGRKQGGYTHMQOQEGDTAGLKESPLQDTEDGSESEPSSETS 64
Db 924 KDEFSV-DKEASAHISGD---SGLSKEFDQEKAN-----DRLDTVLER-SEEHADSK 973
Qy 65 DAKSTPTAEDVTAPLVDEGAPGKQAAAQPHTEIPEGTTAEAA--GIGDTSPLEDEAAGHV 122
Db 974 HAKTEAGD-EIETFGVLYEQALAK-DLSIPTDASSEKAEGKGLSSVPEI----- 1023
Qy 123 TQPESEKVVQEGF-----LREPGLSHQLMSGM-----PGAPLLPEGPREA 166
Db 1024 -AEVPSKVEOGLDFAVQGLDVKISDFG-----QMASGLNIDRRATELKLQATQDM 1076
Qy 167 TROPSTGTPEDT-----EGGRHAPPELLKHQLGDD--LHQEGPPLK---GAGGK 209
Db 1077 T--PSSKAPQADAFMGVESHMKEGTKVSETEVKQKAPDLVHQEAVDKVEESYSSGE 1134
Qy 210 ERPGSKEEVEDRDVDSESSPQDS-----PP--SKASPAQDGRP----- 245
Db 1135 HESLTMESLKADEGKETSPESLIQDEIAVKLSVEIPCPVAVSEADLATDRADVQMEF 1194
Qy 246 ---PQTAAREATSPGPAEAGIPLVPDLFSLKVSVEIPASEPDGPGSVGRAKGQDAPLEFT 302
Db 1195 IQPKESKETPDITSPDVAEPLHETIVSE-PAEQSEEEE-----IEAQGEYDKLLFR 1249
Qy 303 F-----HVEITPNVQK-----EQAHSEHGLGRAAFP 329
Db 1250 SDTLQITDLGVSAREFVETCPSEHKGVIESTIEDDFTVVQVTTTDEGESGSHSVRF 1309
Qy 330 APGEGPEARGPSLGEDTKADLPPESEKQ-----PAAAPRGKPVSRVPOLKARVMS 380
Db 1310 AALQEQEVERRPSPHDEEVEEAEAAQAEKPDGSPAPASPEREEVA-LSEYKTYTD 1368
Qy 381 KSKDGTGSDD-----LSPKLTGSSDPL-----KKAKTSTRSS-AKTLKNRP 408
Db 1369 DYKDETTIDSDIMDADSLWDTQDDRSIMTEQLETPKEEKAKEARRSSLEKHKRKP 1428
Qy 409 C-----LSPKLTGSSDPL-----IQPSSPA--VCPEPP 436
Db 1429 FKTGRGRISTPERKVAKEPSTVSRDEVRKRAVYKAEAKKTEVQAHSPSRKFLKPA 1488
Qy 437 ---SSPKHVSVTSRTGSSGA-----KEMKLGADGKTKIATPRGAAP-----PG 478
Db 1489 IKYTRPHTLSCVKKRTTAAGGESALAPSVFKQAKDKVSDGVTSPKRSRLPRSSILPP 1548
Qy 479 OKG-----QANATRIPAKTP-----PAPKTPPSS---GEPPKSG 509
Db 1549 RRGVSGDRDENSEFSLNSSISSARRTTRSEPIRRAGKSGTSTPTTGTATPTGTPPSS 1608
Qy 510 DRGYSYSPGSGTSGSRTPSLTPPP-----TREPKVAVVTRTPPKSPSAKSLRQTAP 564
Db 1609 SR---TPGTGTP-SYRTPHTPGTKSAILVPSEKKVAIIRTPPKSPATPK-QLRLIN 1662
Qy 565 VPMPLKKNVSKIGSTENLKHQGGGVQIINKKLDLSNVQSKGSKDNIKHVPVGGG 624
Db 1663 QPLPDLKKNVSKIGSTDNIKYQPKGGQVRLINKKIDFSKVQRCGSKDNIKHAGGNNVQ 1722
Qy 625 IVYKVPDLKSVTSKCGSLGNIIHKPGQGVQVSEKLDKDRVQSKIGSLDNIHVPVGGG 684
Db 1723 IVTKKIDLSHVTSCGSLANIRHRRGGVRKTESVKLDFEKAQAQKVGSLDNAHVPGG 1782
Qy 685 NKKIETHKLTFRENAKAKTDHGAIEIVYKSPVVGDTSPRHLNSVSTGSDIMVDSPLQAT 744
Db 1783 NVKIDSKLNFREHAKARVDHGAIEITQPGRSVASPRRLSNVSSGSLNLESPLQAT 1842
Qy 745 LADEVASLAKOGL 758
Db 1843 LAEDVTAALAKOGL 1856
```


Db	924	KDQSV-DREASAHISGDK---SGLSKEDFDQKKAN-----DRJDTVLEK-SBHDASKE	973
Qy	65	DAKSTTAEDVTAPLVDGAPGKQAAAPHTIEPIGTTAEBA--GIGDTPTSLEDAAGHV	122
Db	974	HAKKTEAGD-EIETFTGLVTVQALAK-DLSIPTDASSEKAENGLSSVPEI-----	1023
Qy	123	TQPPESGKVUVQEGF-----LRPGPPGLSHQLMSCM-----PGAPLLPEGPREA	166
Db	1024	-AEVPSKKVEOGLDFAVOGQLDVAKISDFG-----QMASGLDIDRRATELKLLEATQDM	1076
Qy	167	TROPSTGTPEDT-----EGGRHAPELLKHOLLGD--LHQEGPPLK--GAGGK	209
Db	1077	T---PSSKAQADAPMGVESHKKEGTVSETEVKQYAKPDLVHQEAVDKEESYESSGE	1134
Qy	210	ERPGSKEEVEDRDVDDESSPQDS-----PP--SKASPAQDGRP-----	245
Db	1135	HESLTMESLKADEGKKEETSPESLSIQDETAVKLVSVEIPCPPAVSEADLATDERADVOMEF	1194
Qy	246	--POTAAAREANTSIPGFPAEGAELPVDVFLSKVSTIEIPASEPDGPGVGRAKGODAPLEFT	302
Db	1195	IQGPKEESKETPDISITPSDVAEPDHEITVSE-PAEIQSEEEE---IEAQGEYDKLULFR	1249
Qy	303	F-----HVEITPNVQK-----EQARSEEHGLGRAAFP	329
Db	1250	SDTLQITDLGVSGAREFVETCPSEHKGVIESVWTIEDFTTVQVTTTDEGESSHVSFR	1309
Qy	330	APGEGFEARGPSLGEDTKADLPSESEKQ-----PAAAPRGKVPVSRVPOLKARMS	380
Db	1310	AALQOEPEERRPSPHDDEEFVEFEAAEAQAEKDGSPAPAPASPEREVA-LSEYKTEYD	1368
Qy	381	KSXDGKTSDD-----KKAKTSTRSS-AKTLKNRP	408
Db	1369	DYKDETTIDDSIMDADSLWVDTQDDRSIMTLEQLETIPEKEAKEARRSLEKHREKP	1428
Qy	409	C-----LSPKLPTPGSSDPL-----IQPSPA--VCPPEP	436
Db	1429	FKTGRGRISTPERKVAKEPSTVSRDEVRRKAVYKKAELAKKTEVOAHSPSRKFIKPA	1488
Qy	437	---SSPKHVSVTSRPGSSGA-----KMKLKGADGKTATPRGAAP-----PG	478
Db	1489	IKYTRTHLSVCKRKTATAGGESALAPSVFKQAKDKVSDGVTKSPEKRSSLPRSSILPP	1548
Qy	479	QKG-----QANATRIPAKTP-----PAKPTTPSS-----GEPKPSG	509
Db	1549	RRGVSGDRDENSFSLNSSISSARRTTRSEPIRRACKSGTSTPTPGSTAITPGTPPSYS	1608
Qy	510	DRSGYSRPGTTPGSRSTPSILPTTP-----TREPKVAVVTRTPKSPSSAKSRLQATP	564
Db	1609	SR-----TPGTGTPG-SYKPTPTGTPKSAILVPEKSKVAIIRTPPKSPATPK-OLRLIN	1662
Qy	565	VPMPDLKNVKSIGSTPENILKHOPGGGKVQIINKKLUDLSNVQSGKSGDNIKHVPGGGSVQ	624
Db	1663	QPLPDLKNVKSIGSTDNIKYQKPGQVQIVTKIDLSH-----	1701
Qy	625	IVYKPVDLKSVTSKCSGLNIIHHKPGGGOVEVYKSEKLFKORVOSKGTGSLDNITHVPGGG	684
Db	1702	-----VTSKCSGLNIIHRHFGGGRVKVTSYKLFDFEKAQAKVGLSDNHAHVPGGG	1751
Qy	685	NKKIETHKLTFFRENAKAKTDHGAETVYKSPVSGDTPSRHLSNVSVSTSGSIDMVDSPOLAT	744
Db	1752	NVKIDSQKLNFRHAKARYDHGAETITQSPGRSVASPRRLSNVSSSGSINLLESQOLAT	1811
Qy	745	LADVASASLAKOGL	758
Db	1812	LAEDVTAALAKOGL	1825

[illegible]

```

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51331
; LENGTH: 1825
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51331

```

Query Match	20.5%	Score 811;	DB 5;	Length 1825;
Best Local Similarity	29.3%	Pred. No. 7.9e-29;		
Matches 285;	Conservative 122;	Mismatches 275;	Indels 292;	Gaps 40;

Qy	5	RQEFVEMDHAGTYGLGDRKRDQGGVYTMHQDQEGDQDAGLKESPLQTPPTDEGSEBPGSETS	64
Qy	924	KDFSV--DKEASAHISGDK---SGLSKPEQDKKAN---DRLDTVLEK-SEHDASKE	973
Qy	65	DAKSTPTAEDVTAPLVDGAPCKOAAOPTHTEIGCTAAEEA--CIGDTPSLEDEAAGHV	122
Db	974	HAKTEEAGD-ETIETGLGVYTEQAUK--DLISPTDASSERAKGLSSVPEI-----	1023
Qy	123	TQPESEKVVQEGF-----LREPGPGLSHQLMSG---PGAPLLPEGPREA	166
Db	1024	-AEVPSKKVEQGLDFAVQQLDVKISDFG-----QMASGLNIDRRATELKLLEATODM	1076
Qy	167	TROPSTGTGPEDT-----EGGRIHAPELLKHOLLGD--LHQEGPPLK---GAGCK	209
Db	1077	T--PSSKAPOEADAFMGVSEGHMKETGYSETVKQKAKPOLVHQEADVKEESYESSGE	1134
Qy	210	ERPGSKEVEDRDVDSESPQDS-----pp--SKASPAQGRP-----	245
Db	1135	HESLWTESLKADEGSKKETSPESLIODETAVKLSVEIPCPPAVSEADLATDERADVQMEF	1194
Qy	246	---PQTAAREATSIPIGPAGGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQODAPLEFT	302
Db	1195	IQGPKEESKETPDISITPSDVAEPPLHETIVSE-PAEIOSEEEE---TEAQGYDKLLFR	1249
Qy	303	F-----HVEITPNVK-----BOAHSEBHLGRAAPPG	329
Db	1250	SDLIQITDLGVSGAREFEVETCSEHKGVIESVVIEDDFTIVVQTTTDEGESGSHSRP	1309
Qy	330	APGEGPEARGPSIGEDTKKADLPPESEKO-----PAAAPRCKPVSVRPQLKARWVS	380
Db	1310	AALEQPEVERRRPSPHDEEFVEBEAAEAQAEPKDGSPEAPASPEREEVA--LSEYKTET	1368
Qy	381	KSDDGTGSDG-----KKAKTSTRSS-AKTLKNRP	408
Db	1369	DYKDDETTIDDSIMDASLSLVDTODDRSITMTQLETIPEKKAERKARSLEKHRRKEP	1428
Qy	409	C-----LSPKLPPTGSSDPL-----IQPSSPA--VCPPEPP	436
Db	1429	EKTGRGRISTPERKVAKEPSTVSRDEVRRKAVYKKAELAKKTEVQAHSRPFILKPA	1488
Qy	437	---SSPKHVSSTVSRGSSGA-----KEMKLGADGKTIXATPGAAP-----PG	478
Db	1489	IKYTRPTHLSCVRKRTTAAGGESALAPSVFKQAKDKVSDGVTKSPKRSLSLPRPSSILPP	1548
Qy	479	QKG-----QANATRIPAKTP-----PAKPTPPSS---GEPKPSG	509
Db	1549	RRGVSGDRDENSELSNSSITSSSARRTIREPIRACKSGTSTPTTPGTATPTGTPPPSY	1608
Qy	510	DRSGYSPPSGPGTPGGRSKRTPSLTPP-----TREPKKVAVVTRTPPKSPSSAKSRLO	564
Db	1609	SR-----TPGTPGTP-SYPRTPHPGTPKSAIILVPEKKVAILIIRTPPKSPATPK-Q	1662
Qy	565	VPMPDLKNVASKTGSTENLKHOPGGCKVOIINKKLIDNLSVQSCGSKDNIKHVPGGSGVQ	624
Db	1663	QPLPDLKNVASKTGSTNDNIKIYDPKGQVOIVTKMKIDLSH-----	1701
Qy	625	IVYKPDLSKVTSCGCSLGNIIHKPGGGQGVYKSEKLDPKDRVQSKIGSLONITHVPGGG	684


```

Db 1702 -----VTSKCSGLNIRHPGGGRVKIESVKLDFEKAQAVGSLDNAHHVPGG 1751
QY 685 NKIETHKLFRENAKTDHGAIEIVKSPVVSDDTSPRHLNVSTGSDMDVDSQLAT 744
Db 1752 NVKIDSKLNFREHAKRVDHGAIEITOSPGRSVASPRRLSNVSSGSLNLESQLAT 1811
QY 745 LADEVASLAKOGL 758
Db 1812 LAEDVTAALAKOGL 1825

RESULT 12
US-09-724-676-51332
; Sequence 51332, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51332
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51332

```

```

Query Match      18.9%; Score 750.5; DB 5; Length 471;
Best Local Similarity 38.3%; Pred. No. 1.1e-26;
Matches 199; Conservative 70; Mismatches 128; Indels 123; Gaps 18;

QY 297 APL-EFTFHVETPNVQKQAHSEHLAGRAA--FP-----GAPGE-GPEARGPSLGEDT 346
Db 17 APLTEASAHSH--PPEIKDQGGAGELVRSANGFPYREDEGAFGEHGSQGTYSNTKENG 74
QY 347 KEADLPPEPEKQPAAPRGPVSRVQLKARMYSKSDGTGDDKAKTSTRSSAKTLKN 406
Db 75 INGEL--TSADRETAEEVSARIVQVVTAEAVALK-----GEQKEAQHKDQTAAL---- 123
QY 407 RPLCLSPKLPFGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGA-----KEMK 458
Db 124 -----PLAAEETANLPPSPPPSPASEQTVTVEEAAGGESALAPSVFQKQAK 168
QY 459 LKGADGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPKRSLSLPRSSILPPRGVSGDRDENSEFSLNSSISSARRTRSEPIRR 228
QY 494 -----PAPKTPPSS-----GEPKSGDRSGYSSPGSGPTGSRSTRPSLTPTP-----TR 538
Db 229 AGKSGTSTPTPGSTAITPGTPPSYSSR---TPGTPGTP-SYPRTPHTPGTPKSAILVP 283
QY 539 EPKKAIVVTRTPPKSPSSAKSRLOTAPVPMPLKNVKSIGSTENLKHQPGGKQVQIINKK 598
Db 284 SEKKVAILRTPPKSPATPK-QURLINQPLDPLKNVKSIGSTDNIKYQKPGQGVQIVTKK 342
QY 599 LDLSNVQSKCGSKDNKIKHVPGGGSVQIVYKPVDSLKVTSCGSLGNIHHKPGGGOVEVKS 658
Db 343 IDLSH-----VTSCGSLKNIRHRRPGGRVKIES 371
QY 659 EKLDKDRVQSKIGSLDNITHVPGGKNKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
Db 372 VKLDFEKAQAKVGLSDNAHHVPGGNGVKIDSQKLNFRHAKARVDHGAIEITQSPGRSS 431
QY 719 DTSRHLNVSTGSDMDVDSQLATLADEVASLAKOGL 758
Db 432 VASPRRLSNVSSGSLNLESQLATLAEDVTAALAKOGL 471

RESULT 13
US-09-724-676A-51332
; Sequence 51332, Application US/09724676A

```

```

; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51332
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51332

Query Match      18.9%; Score 750.5; DB 5; Length 471;
Best Local Similarity 38.3%; Pred. No. 1.1e-26;
Matches 199; Conservative 70; Mismatches 128; Indels 123; Gaps 18;

QY 297 APL-EFTFHVETPNVQKQAHSEHLAGRAA--FP-----GAPGE-GPEARGPSLGEDT 346
Db 17 APLTEASAHSH--PPEIKDQGGAGELVRSANGFPYREDEGAFGEHGSQGTYSNTKENG 74
QY 347 KEADLPPEPEKQPAAPRGPVSRVQLKARMYSKSDGTGDDKAKTSTRSSAKTLKN 406
Db 75 INGEL--TSADRETAEEVSARIVQVVTAEAVALK-----GEQKEAQHKDQTAAL---- 123
QY 407 RPLCLSPKLPFGSSDPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGA-----KEMK 458
Db 124 -----PLAAEETANLPPSPPPSPASEQTVTVEEAAGGESALAPSVFQKQAK 168
QY 459 LKGADGKTKIATPRGAAP-----PGQKG-----QANATRIPAKTP--- 493
Db 169 DKVSDGVTKSPKRSLSLPRSSILPPRGVSGDRDENSEFSLNSSISSARRTRSEPIRR 228
QY 494 -----PAPKTPPSS-----GEPKSGDRSGYSSPGSGPTGSRSTRPSLTPTP-----TR 538
Db 229 AGKSGTSTPTPGSTAITPGTPPSYSSR---TPGTPGTP-SYPRTPHTPGTPKSAILVP 283
QY 539 EPKKAIVVTRTPPKSPSSAKSRLOTAPVPMPLKNVKSIGSTENLKHQPGGKQVQIINKK 598
Db 284 SEKKVAILRTPPKSPATPK-QURLINQPLDPLKNVKSIGSTDNIKYQKPGQGVQIVTKK 342
QY 599 LDLSNVQSKCGSKDNKIKHVPGGGSVQIVYKPVDSLKVTSCGSLGNIHHKPGGGOVEVKS 658
Db 343 IDLSH-----VTSCGSLKNIRHRRPGGRVKIES 371
QY 659 EKLDKDRVQSKIGSLDNITHVPGGKNKKIETHKLTFRENAKAKTDHGAIEIVKSPVVS 718
Db 372 VKLDFEKAQAKVGLSDNAHHVPGGNGVKIDSQKLNFRHAKARVDHGAIEITQSPGRSS 431
QY 719 DTSRHLNVSTGSDMDVDSQLATLADEVASLAKOGL 758
Db 432 VASPRRLSNVSSGSLNLESQLATLAEDVTAALAKOGL 471

RESULT 14
US-10-070-611-1
; Sequence 1, Application US/10070611
; GENERAL INFORMATION:
; APPLICANT: VON BERGEN, Martin
; APPLICANT: BIERNAT, Jacek
; APPLICANT: MANDELKOW, Eva-Maria
; APPLICANT: MANDELKOW, Eckhard
; TITLE OF INVENTION: A MINIMAL TAU PEPTIDE FOR THE NUCLEATION OF PATRED HELICAL FRA
; FILE REFERENCE: 029976/0101
; CURRENT APPLICATION NUMBER: US/10/070,611
; CURRENT FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: PCT/EP00/08863
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: EP 99117805.4
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1

```


; SEQ ID NO 1
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-070-611-1

Query Match 16.6%; Score 657; DB 6; Length 125;
Best Local Similarity 100.0%; Pred. No. 4.4e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 561 QTAPVMPDLKNNVSKIGSTENLKHOPGGKVOIINKKLDLSNVQSKGSKDNIKHVPGG 620
Db 1 QTAPVMPDLKNNVSKIGSTENLKHOPGGKVOIINKKLDLSNVQSKGSKDNIKHVPGG 60
QY 621 GSVQIVYKPVLDLKVTSKCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHV 680
Db 61 GSVQIVYKPVLDLKVTSKCGSLGNIHHKPGGGQVEVSEKLDKDRVQSKIGSLDNITHV 120

QY 681 PGGGN 685
Db 121 PGGGN 125

RESULT 15

US-10-100-957A-152
; Sequence 152, Application US/10100957A
; GENERAL INFORMATION:
; APPLICANT: Kapur, Ravi
; APPLICANT: Giuliano, Kenneth A.
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1A
; CURRENT APPLICATION NUMBER: US/10/100,957A
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 152
; LENGTH: 1125
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-100-957A-152

Query Match 15.7%; Score 621.5; DB 6; Length 1125;
Best Local Similarity 28.5%; Pred. No. 1.6e-20;
Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;

QY 43 LKESPLQTPEDGS---EE--PGSETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTE 96
Db 317 LPTPEDESKDVAPPMEEIIVCGNNTTSPKETET----TLPIKMDLAPEDVLLTKETE 372
QY 97 IPEGTTAAEAGIGDTPSLEDEAAGHVTVQPESGKVVOEGFLREPGPPGLSHQIMSGMP-- 154
Db 373 L-----APAKGWSVLSIEIEALAKNDVSAEI--PVAQETWVSEVVLATEVVLPSDPIT 426
QY 155 -----GAPLLPGPREATRQPGTGPEDTEGGRHAPPELLKHOLLGDHLQEGPPLKAG- 207
Db 427 TLTKDVTLPLEAERPLVTMTPS-LETEMTLGKETAPPTETN--LG-MAKDMSPLESEV 482
QY 208 --GKE-----RPGSKEEVEDRDVDDESSPDSPSK----- 236
Db 483 TLGKDVVILPETKVAEFNNVTPLSEEVTSVKDMSPSAETAPLAKNADLHSGTELIVDN 542
QY 237 -ASPAQD-----GR---PPQTAAREAT 254
Db 543 SMAPASDLALPLETKVATVPDKGTVQTEEKPREDSQLASMQHKGQSTVPCTASPE-- 600
QY 255 SIFGFPAGGAIPLPVDFLSKVST----EIPASEPDGSPVGRAGQDAPLETFHVEITFN 310
Db 601 --PVKAAEQMSTLPIDAPSPLENLEQKETPGSQSPSCGVSQEEAKA-----VGVTCN 654
QY 311 VQKEQAHSEHLGRAAFPAGPGGPEARG-----PSLGEDTKADLPEPSEKQPA--- 360
Db 655 DITTPPNKE-----PPPSPEKKAKPLATTQTPAKTSTSKAKTQTPSLPKQAPPT 703

QY 361 -----AAPRGKP-----VSRVPQLKARVMVSKSDGTGSD--DKKAKTS 396
Db 704 SGGLNKKPMSLASGVPAAHPKRPAAATATATPSTLPARDV-KPKPITEAKVAEKRTSPS 762
QY 397 TRSSAKTLKNRPCLSKLPPTPGSSDPLIQ--PS--SPA-VCPEPPSS-----PKHVSSV 445
Db 763 KPSSAPALKPGPKTPTTVSKATSPSTLVSTGPPSRSPATTLPKRPTSLKTEGKPADVKRM 822
QY 446 TSRTGSSGAKEMKLGADGKTKIATPRGAAPPQKQGOANATRIPAKTPPAPKTPPSSGEP 505
Db 823 TAKSASADLSRSKTTSSASSVKRNTTPTGAAPP--AGMTSTRV-----KPMAPSRSRG 873
QY 506 PKSGDRSGYSSPGSPGTPGSRSTPSPSLPTPTPREPKKVAVVVRTPPKPSSAKSLQTA 565
Db 874 ALSVDKK-----PT-----STKPSSSAPRVSRLATT-V 900
QY 566 PMPDLKNVSKIGSTENLKHOPGG----- 589
Db 901 SAPDLKSVRSKVGSTENIKHQPGGGRAKVEKTEAATTAGKPEPNAVTKAAGSTASAKP 960
QY 590 --GKVOIINKKLDLSNVQSKGSKDNIKHVPGGGSVQIVYKPVLDLKVTSKCGSLGNIHH 647
Db 961 PAGKVQIVSKVSYSHIQSKVCSDNKHVPGCGNVQIQNKVKVDISKVSSKCGSKANIKH 1020
QY 648 KPGGGQVEVSEKLDKDRVQSKIGSLDNITHVPGGGNKKIE 689
Db 1021 KPGGGDVKIESOKLNFKEKAQAKVGSGLDNVGHFPAGGAVKTE 1062

Search completed: January 28, 2003, 14:39:20
Job time : 30.1541 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 ; Search time 15.7356 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-5

Perfect score: 3965

Sequence: 1 MAEPQPEVMEHDHAGTYGL.....SPQLATLADEVSASLAKOGL 758

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116.5	53.4	441	4	US-08-244-603A-1
2	2077.5	52.4	441	4	US-08-913-915-5
3	1776.5	44.8	383	4	US-09-336-038-4
4	1590	40.1	352	2	US-08-726-306A-17
5	1590	40.1	391	2	US-08-244-951A-10
6	1590	40.1	391	2	US-08-389-011-23
7	1590	40.1	391	3	US-08-403-917A-23
8	1590	40.1	391	4	US-08-348-952A-23
9	1586	40.0	351	1	US-08-159-969-2
10	733	18.5	140	4	US-08-913-915-9
11	621.5	15.7	1125	4	US-09-513-783A-152
12	621.5	15.7	1610	4	US-09-513-783A-22
13	593	15.0	112	3	US-08-666-360-1
14	575	14.5	109	4	US-08-913-915-3
15	563	14.2	106	3	US-08-778-404B-1
16	561	14.1	108	4	US-08-913-915-2
17	546.5	13.8	109	4	US-08-913-915-1
18	546	13.8	140	4	US-08-913-915-10
19	416	10.5	160	2	US-08-726-306A-35
20	364	9.2	67	2	US-08-389-011-1
21	364	9.2	67	3	US-08-403-917A-1
22	364	9.2	67	4	US-09-348-952A-1
23	339	8.5	67	2	US-08-244-951A-1
24	283.5	7.2	1057	3	US-08-931-820-4
25	283	7.1	1185	4	US-09-041-886-23
26	280.5	7.1	1442	2	US-08-316-650-12
27	280.5	7.1	1442	5	PCT-US95-02251-12

28	278.5	7.0	1078	3	US-08-963-825-21	Sequence 21, Appl
29	278.5	7.0	1078	4	US-09-500-811-21	Sequence 21, Appl
30	278.5	7.0	1078	4	US-09-570-573-21	Sequence 21, Appl
31	278.5	7.0	1078	4	US-09-548-608-21	Sequence 21, Appl
32	274	6.9	941	4	US-07-757-022B-14	Sequence 84, Appl
33	274	6.9	1022	4	US-07-757-022B-84	Sequence 74, Appl
34	274	6.9	1038	4	US-07-757-022B-74	Sequence 58, Appl
35	274	6.9	1049	4	US-07-757-022B-58	Sequence 104, App
36	274	6.9	1140	4	US-07-757-022B-104	Sequence 44, Appl
37	274	6.9	1270	4	US-07-757-022B-44	Sequence 42, Appl
38	274	6.9	1311	4	US-07-757-022B-42	Sequence 142, App
39	274	6.9	1313	4	US-07-757-022B-142	Sequence 50, Appl
40	274	6.9	1314	4	US-07-757-022B-50	Sequence 46, Appl
41	274	6.9	1320	4	US-07-757-022B-46	Sequence 60, Appl
42	274	6.9	1320	4	US-07-757-022B-60	Sequence 48, Appl
43	274	6.9	1354	4	US-07-757-022B-48	Sequence 40, Appl
44	274	6.9	1361	4	US-07-757-022B-40	Sequence 52, Appl
45	274	6.9	1363	4	US-07-757-022B-52	

ALIGNMENTS

RESULT 1

US-08-244-603A-1

; Sequence 1, Application US/08244603A

; Patent No. 6200768

; GENERAL INFORMATION:

; APPLICANT: Mandelkow, Eva-Maria

; APPLICANT: Mandelkow, Eckhard

; APPLICANT: Lichtenberg-Kraag, Birgit

; APPLICANT: Biernat, Jacek

; APPLICANT: Drewes, Gerard

; APPLICANT: Steiner, Barbara

; TITLE OF INVENTION: No. 6200768el Tools For The Diagnosis And

; TITLE OF INVENTION: Treatment Of Alzheimer's Disease

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

; ADDRESSEE: Borun

; STREET: 233 South Waker Drive, 6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Tape

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/244,603A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Joseph A. Williams, Jr.

; REGISTRATION NUMBER: 38,659

; REFERENCE/DOCKET NUMBER: 28384/32778

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-484-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 441 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-244-603A-1

Query Match 53.4%; Score 2116.5; DB 4; Length 441;

Best Local Similarity 58.2%; Pred. No. 1.3e-133;

QY 361 AAPRGKPVSRVPOQLKARWYSKSDGTGSDDDKAKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420
Db 67 -----ARMVSKSDGTGSDDKA----- 84
QY 421 DPLQPSPPAVCPPEPPSPKHSVTSRTGSSGAKEMKLGADGKTKIATPRCAAPPGOK 480
Db 85 -----KGADGKTKIATPRCAAPPGOK 105
QY 481 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 540
Db 106 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 165
QY 541 KKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSIGSTENLKHQPGGKVOIINKLKD 600
Db 166 KKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSIGSTENLKHQPGGK----- 216
QY 601 LSNVQSKGSKDNIKHVPGGGVOIVYKVDLSKVTSKGSLGNIHHKPGGQVEVKSEK 660
Db 217 -----VOIVYKVDLSKVTSKGSLGNIHHKPGGQVEVKSEK 254
QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 720
Db 255 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 314
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 315 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 352

RESULT 5

US-08-244-951A-10
; Sequence 10, Application US/08244951A
; Patent No. 5843779

GENERAL INFORMATION:

; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN; VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016

COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/244, 951A
; FILING DATE: 19-JAN-1995

CLASSIFICATION:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 10-DEC-1993

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992

ATTORNEY/AGENT INFORMATION:

; NAME: CHARLES A. MUSERLIAN

; REGISTRATION NUMBER: 19,683

; REFERENCE/DOCKET NUMBER: 410.003A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000

; TELEFAX: (212) 661-8002

; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; FEATURE:
; NAME/KEY: mTHMPH- τ fusion protein
US-08-244-951A-10

Query Match

Best Local Similarity 40.1%; Score 1590; DB 2; Length 391;

Matches 352; Conservative 46.4%; Pred. No. 1.4e-98;

Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPQREFEVMEHDHAGTYGLGRKDDQGGYTMHQDQEGDTDAGLKESPLQPTFEDGSEEPG 60
Db 40 MAEPQREFEVMEHDHAGTYGLGRKDDQGGYTMHQDQEGDTDAGLK----- 83
QY 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAAPHITEIPEGTTAEAGIGDTPSLDEAAG 120
Db 84 -----AEEAGIGDTPSLDEAAG 101
QY 121 HVTQBPESGKVVQEGFLREPGLSHQMSGMPGAPLLPEGPREATRQPSGTGPEDETEG 180
Db 102 HVTQ----- 105
QY 181 GRHAPELLKHQLLDLHQLGDLHQLGAGGKERPGSKKEEVEDRDVDDESSPDSPSKASPA 240
Db 106 ----- 105
QY 241 QDGRPPQTAAREATSIPIGFPAGGAIPLVDFLSKVSTEIPASEPDGPSVGRAKQDAPLE 300
Db 106 ----- 105
QY 301 FTFHVEITPNVQEQAHSEHGLRAAFPAPGEGPEARGPSLGEDTKEADLPEPSEKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVPOQLKARWYSKSDGTGSDDKAKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420
Db 106 -----ARMVSKSDGTGSDDKKA----- 123
QY 421 DPLQPSPPAVCPPEPPSPKHSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPGOK 480
Db 124 -----KGADGKTKIATPRGAAPPGOK 144
QY 481 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 540
Db 145 GOANATRIAPAKTPPSSGEPKSGDRSGYSSPGSGTSGSRSTPSLPTPTREP 204
QY 541 KKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSIGSTENLKHQPGGKVOIINKLKD 600
Db 205 KKVAVRTPPKSPSSAKSRLOTPAPVMPDLKNVKSIGSTENLKHQPGGK----- 255
QY 601 LSNVQSKGSKDNIKHVPGGGVOIVYKVDLSKVTSKGSLGNIHHKPGGQVEVKSEK 660
Db 256 -----VOIVYKVDLSKVTSKGSLGNIHHKPGGQVEVKSEK 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAEIYKSPVVSQGT 353
QY 721 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 758
Db 354 SPRHLSNVSTGSDIMVDSPOLATLADEVASLAKOGL 391

RESULT 6

US-08-389-011-23

; Sequence 23, Application US/08389011

; Patent No. 5861257

GENERAL INFORMATION:

; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;

; APPLICANT: VANMECHELEN, EUGEN; VAN DE VOORDE, ANDRE

; TITLE OF INVENTION: MONOCLONAL ANTIBODIES


```

; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/389,011
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/403,917
; FILING DATE: 19-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/403,916
; FILING DATE: 19-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,951
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 10-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410,003-1-CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
;
US-08-389-011-23
Query Match 40.1%; Score 1590; DB 2; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.4e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

QY 1 MAEPROEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTHAGLKEGSPLOTPTEDGSEEPG 60
Db 1 |||||
QY 40 MAEPROEFVEMDHAGTYGLGDRKQGGYTMHQDQEGDTHAGLKEGSPLOTPTEDGSEEPG 83
Db 40 |||||
QY 61 SETSDAKSTPTAEDYAPLVDEGAPGKQNAAPHTPEIGTAAEAGIGDTPSLDEAAG 120
Db 61 |||||
QY 121 HVTQPEGKVVQEGFLREPPGPPGLSHQLMSGMPGAPLLPEGPREATQPSCTGPDTEG 180
Db 121 |||||
QY 102 HVTQ----- 105
Db 102 HVTQ----- 105
QY 181 GRHAPELLKHQLLDLHQLHQPPLKAGGKERPGSKEEVEDRDVDDESSPDSPSKASPA 240
Db 181 ----- 105
QY 241 QDGRPPTAAREATSIPIGPAEATPLPVDLFLSKVSTEIPASEPDGSPVGRAGKODAPLE 300
Db 241 ----- 105

```

```

Db 106 ----- 105
QY 301 FTFHVEITPNVQEQAHSEHGLRAAFPCAPFEGPEARGPSLGEDTKADLPPESEKQPA 360
Db 106 ----- 105
QY 361 AAPRGKPVSRVQPKARMVSKSDGTGSDDKKAKKTSTRSSAKTLKNRPCLSPKLPPTPGSS 420
Db 106 ----- 123
QY 421 DPLIQSPSPAVCPPEPPSPKHVSSVTSRTGSSGAKEMKLKGADGKTATPRGAAPGQK 480
Db 124 ----- 144
QY 481 QOANATRIPTAKTPPAKTPPSSGEPKSGDRSGYSSPSGCTPGSRSTRPSLPTPTPREP 540
Db 145 QOANATRIPTAKTPPAKTPPSSGEPKSGDRSGYSSPSGCTPGSRSTRPSLPTPTPREP 204
QY 541 KKVAVVRTPPKSPSSAKSRLQTAQVPMPLDNVSKIGSTENLKHQPGGKGVQIINKLD 600
Db 205 KKVAVVRTPPKSPSSAKSRLQTAQVPMPLDNVSKIGSTENLKHQPGGK----- 255
QY 601 LSNVQSKGSKDNIKHVPGGSVQIVYKPVDSLKVTSCGSLGNHHPKGGQGVKSEK 660
Db 256 ----- 293
QY 661 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGNNKKTETHKLTFRENAKAKTDHGAEIVYKSPVWSGDT 353
QY 721 SPRHLSNVSTGSDIMVDPOLATLADDEVASLAKOGL 758
Db 354 SPRHLSNVSTGSDIMVDPOLATLADDEVASLAKOGL 391

RESULT 7
US-08-403-917A-23
; Sequence 23, Application US/08403917A
; Patent No. 6010913
; GENERAL INFORMATION:
; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN;
; APPLICANT: VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,917A
; FILING DATE: 19-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/256,167
; FILING DATE: 27-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,951
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 10-DEC-1993

```



```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.003-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
;
US-08-403-917A-23

Query Match 40.1%; Score 1590; DB 3; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.4e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

Qy 1 MAEPQFEFVMEHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPTEGSEEPG 60
Db 40 MAEPQFEFVMEHAGTYGLGDRKQGGYTMHQDQGGDTDAGLK-----83
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 84 -----AEEAGIGDTPSLEDEAAG 101
Qy 121 HVTQEPESGKVVOEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
Db 102 HVTQ-----105
Qy 181 GRHAPELLKHQLLDLHQEGPPLKAGGKERPGSKEVEVDROVDRESSPODSPSKASPA 240
Db 106 -----105
Qy 241 QDGRPPQTAAREATSIPGFPAEGAIPLPVDFLSKVSTEIPASEPDGPDGVS CRAKQDAPLE 300
Db 106 -----105
Qy 301 FTFHVEITPNVORQAHESEHLGRAAPGAPGEGPEARGPSLGEDTKEADLPEPSEKQPA 360
Db 106 -----105
Qy 361 AAPRGKPVSRVPQKLKRWVSKSDGTGSDDKKAKTSTRSSAKTLKNRPCLSPKLTPPGSS 420
Db 106 -----123
Qy 421 DPLIQSSPAVCPPEPPSSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPPQOK 480
Db 124 -----KGADGKTKIATPRGAAPPQOK 144
Qy 481 QGANATRIKTPAPKTPPPSGEPKPSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 540
Db 145 QGANATRIKTPAPKTPPPSGEPKPSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP 204
Qy 541 KKVAVRTPKPSASRSRLQTPAPVMPDILKNVSKSIGSTENLKHQPGGKQVQIINKLKD 600
Db 205 KKVAVRTPKPSASRSRLQTPAPVMPDILKNVSKSIGSTENLKHQPGGK-----255
Qy 601 LSNVQSKGSKDNKHPVGGSGVQIVKPYDLSKVTSSKCSLNIHKKPGGGQVEVKSEK 660
Db 256 -----VQIVKPYDLSKVTSSKCSLNIHKKPGGGQVEVKSEK 293
Qy 661 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVKSPVVSQGT 720
Db 294 LDFKDRVQSKIGSLDNITHVPGGKNKIETHKLTFRENAKAKTDHGAIEIVKSPVVSQGT 353
Qy 721 SPRHLNSVSTGSDIMVDSFQLATLADSVASLAKQGL 758
Db 354 SPRHLNSVSTGSDIMVDSFQLATLADSVASLAKQGL 391
```

```
RESULT 8
US-09-348-952A-23
; Sequence 23, Application US/09348952A
; Patent No. 6232437
; GENERAL INFORMATION:
; APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
; APPLICANT: VANMECHELEN, EUGEN;
; APPLICANT: VAN DE VOORDE, ANDRE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
; TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SECRETING THESE
; TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/348, 952A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/403,917
; FILING DATE: 19-JAN-1995
; APPLICATION NUMBER: 08/256,167
; FILING DATE: 27-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/244,951
; FILING DATE: 13-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03499
; FILING DATE: 10-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP/92/403403.6
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 410.003-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
;
US-09-348-952A-23
```

```
Query Match 40.1%; Score 1590; DB 4; Length 391;
Best Local Similarity 46.4%; Pred. No. 1.4e-98;
Matches 352; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

Qy 1 MAEPQFEFVMEHAGTYGLGDRKQGGYTMHQDQGGDTDAGLKESPLQTPTEGSEEPG 60
Db 40 MAEPQFEFVMEHAGTYGLGDRKQGGYTMHQDQGGDTDAGLK-----83
Qy 61 SETSDAKSTPTAEDVTAPLVDEGAPGKQAAQPHTEIPEGTTAEAGIGDTPSLEDEAAG 120
Db 84 -----AEEAGIGDTPSLEDEAAG 101
Qy 121 HVTQEPESGKVVOEGFLREPGLSHQLMSGMPGAPLLPEGPREATRQPSGTGPEDTEG 180
```



```

; NAME: Misrock, S.Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 4697-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-159-969-2

```

Query Match 40.0%; Score 1586; DB 1; Length 351;
 Best Local Similarity 46.4%; Pred. No. 2.2e-98;
 Matches 351; Conservative 0; Mismatches 0; Indels 406; Gaps 4;

Db	102	HVTO-----	105
Qy	181	GRHAPELLKQLLDLHOGEPPLKAGGKERPGSKEEVEDRDVDESSPODSPPSKASPA	240
Db	106	-----	105
Qy	241	QDGRPPOTAAREATSIFGPPAEGAIPLPVDFLSKVSTEIPASEPDGPGSVGRAKQDAPLE	300
Db	106	-----	105
Qy	301	FTFHVETPNVQEQAHSEHLGRAAFPGAPGEPGARGPSLGEDTKEADLPPESEKQPA	360
Db	106	-----	105
Qy	361	AAPRGKPVSRVLPQLKARWVSKSDGTGDDKKAKTSTRSSAKTLKRNPCLSPKLPTPGSS	420
Db	106	-----	123
Qy	421	DPLIQSPSPAVCPPEPPSPKHVSVTSRTGSSGAKEMKLGADGKTKIATPRGAAPGQK	480
Db	124	-----	144
Qy	481	GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP	540
Db	145	GOANATRIPAKTPAPKTPPSSGEPKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTREP	204
Qy	541	KKVAVVTRTPPKSPSSAKSRLQTPVPMPLKKNVSKIGSTENLKHQPGGKQVQIINKKLD	600
Db	205	KKVAVVTRTPPKSPSSAKSRLQTPVPMPLKKNVSKIGSTENLKHQPGGKQVQIINKKLD	255
Qy	601	LSNVQSKCGSKDNKHKVPGGGSQIVYKPVLDLSKVTSSKCSLGNIIHHPGGQGVVEKSEK	660
Db	256	-----	293
Qy	661	LDFKDRVQSKIGSLDNTHVPGGNNKIIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT	720
Db	294	LDFKDRVQSKIGSLDNTHVPGGNNKIIETHKLTFRENAKAKTDHGAIEIVYKSPVVSQGT	353
Qy	721	SPRHLSNVSSSTGSDMVDSFQLATLADEVASLAKQGL 758	
Db	354	SPRHLSNVSSSTGSDMVDSFQLATLADEVASLAKQGL 391	

```

RESULT 9
US-08-159-969-2
; Sequence 2, Application us/08159969
; Patent No. 5492812
; GENERAL INFORMATION:
; APPLICANT: Voorheils, Paul H.
; TITLE OF INVENTION: Diagnostic Method for Alzheimer's
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,969
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/738,778
; FILING DATE: 01-AUG-1991
; ATTORNEY/AGENT INFORMATION:

```

RESULT 10

US-09-913-915-9
 ; Sequence 9, Application US/08913915
 ; Patent No. 6376205

; GENERAL INFORMATION:
 ; APPLICANT:

; TITLE OF INVENTION: INHIBITION OF TAU-TAU-ASSOCIATION

; NUMBER OF SEQUENCES: 12

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/913,915

; FILING DATE:

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 140 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-913-915-9

Query Match 18.5%; Score 733; DB 4; Length 140;

Best Local Similarity 100.0%; Pred. No. 4.8e-42;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 569 DLKNNKIGSTENLKHQPGGKVIINKKLDLSNVQSKGSKDNKIHVPGGSGVQIVYK 628

DB 1 DLKNNKIGSTENLKHQPGGKVIINKKLDLSNVQSKGSKDNKIHVPGGSGVQIVYK 60

QY 629 PVDLSKVTSCGSLGNTHHKPGGQVEVKSLEKDFKDRVQSKIGSLDNITHVPGGNGKKI 688

DB 61 PVDLSKVTSCGSLGNTHHKPGGQVEVKSLEKDFKDRVQSKIGSLDNITHVPGGNGKKI 120

QY 689 ETHKLTFRNKAKTDHGAE 708

DB 121 ETHKLTFRNKAKTDHGAE 140

RESULT 11

US-09-513-783A-152

; Sequence 152, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 152

; LENGTH: 1125

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-513-783A-152

Query Match

Best Local Similarity 15.7%; Score 621.5; DB 4; Length 1125;

Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;

QY 43 LKESPLQTPTEDGS-----EE--PCSEYSDAKSTPTAEDYVDEGAPGKQAAQPHTE 96

DB 317 LPTPEDESTKDVAPPMEETVPGNDTSPKETET---TLPIKMDLAPPEDVLLTKETE 372

QY 97 IPEGTTAEAGIGDTPTSLDEAAGHVTOEPESGKVGQEGFLREPGLSHQLMSGMP-- 154

DB 373 L-----APAKGMVLSLEIEEALAKNDVRSAL-I-PVAQETVVSETEVVLATVVLPSDPIT 426

QY 155 -----GAPLLPEGPREATRQPSGTGPTDEGGRHAPPELLKHQLLDLHOGEPPLKAG- 207

DB 427 TLTGKDVTLPLEAERPLVTOMTPS-LETEMTLGKETAPPTETN--LC-MAKDMSPLESEV 482

QY 208 --GKE-----RPGSKEEVEDRDVDESSQDPSPPSK----- 236

DB 483 TLGKDVVILPETKVAEFNNVTPLSEBEVTSVKDMSPSAETEAFLAKNADLHSGTBLVDN 542

QY 237 -ASPAOD-----PPSPPEKAKPLATTQPAKTSTSKAKTQPTSLPKQAPATT 254

DB 543 SWAPASDLALPLETKVATVPIKDKGTQVTEKEPREDSQLASMQHKGQSTVPCTPASPE-- 600

QY 255 SIPIGPPAEGAIPLPVDLFSKVT-----EIPASEPDPGSPVGRAKGDADPLEFFTHVEITPN 310

DB 601 --PVKAAEQMSTLPIDAPSPLENLEKETPGSQSPGPCSGVSRQBEAKAA-----VGVTGN 654

QY 311 VQKEQAHSEEHGLGRAAFPAGGEPARG-----PSLGEDTKADLPEPESEKOPA--- 360

DB 655 DITTPPNKE-----PPSPPEKAKPLATTQPAKTSTSKAKTQPTSLPKQAPATT 703

QY 361 -----AAPRGKP-----VSRVPQLKARMVSKSKDGTGSD--DKKAKTS 396

DB 704 SGGLNKKPMKSLASGSPVPAAPHKRPAAATATARPSTLPARDV-KPKPITEAKVAERTSPS 762

QY 397 TRSSAKTLNRCPLSPKLTTPGSSDPLIQ--PS--SPA-VCPEPPSS-----PKHVSIV 445

DB 763 KPSSAPALPKPGPKTPTVSKATSPSTLVSTGSPSRSPATTLPKRPTSIKTEGKPADVKRM 822

QY 446 TSRTSGGSAKEKMLKAGDKTKTIATPRGAAPGQKQOANATRIPAKTPAPKTPPSSGEP 505

DB 823 TAKSADISRSKTTTSASSVKRNTTPTGAAP--AGMTSTRV-----KPMASPSRSSG 873

QY 506 PKSGDRSGYSSPGSPGTPGSRSTRPSLPTPTPREPKKVAVVRTPPKSPSSAKSRQLTAPV 565

DB 874 ALSVDKK-----PT-----STKPPSSAPRYSRLATT-V 900

QY 566 PMDLKNVSKIGSTENLKHQPGG----- 589

DB 901 SAPDLKSVRSKVGSTENIKHQGGGRKVERKTEAATTAGKPEPNAVTKAAGSIASAQKP 960

QY 590 --GKVOIINKKLDLSNVQSKGSKDNKIHVPGGSGVQIVYKVDLSKVTSCGSLGNTHH 647

DB 961 PAGKVQIVSKKVSYSHIQKCVSKDNKIHVPGGNGVQIONKKVDSKSSKCGSKANIKH 1020

QY 648 KPGGGQVEYKSEKLDKDRVQSKIGSLDNITHVPGGNGKKIE 689

DB 1021 KPGGGDVKIESQKLNPKKAKAQAQKVGSLDNVGHFEPAGGAVKTE 1062

RESULT 12

US-09-513-783A-22

; Sequence 22, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; APPLICANT: Kapur, Ravi

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 22

; LENGTH: 1610

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: EYFP-DEVD-MAP4-BEFP Construct

US-09-513-783A-22

Query Match

Best Local Similarity 15.7%; Score 621.5; DB 4; Length 1610;

Matches 234; Conservative 91; Mismatches 246; Indels 251; Gaps 33;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:08:34 ; Search time 4.03824 Seconds
(without alignments)
3666.126 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDPVQGIN.....STKTGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	1 DSHUCZ	superoxide dismuta
2	807	98.3	179	4 T43640	superoxide dismuta
3	684.5	83.4	152	2 S36108	superoxide dismuta
4	680	82.8	154	2 J00915	superoxide dismuta
5	678	82.6	154	2 J01192	superoxide dismuta
6	667	81.2	154	1 DSHOCZ	superoxide dismuta
7	659.5	80.3	153	2 S33162	superoxide dismuta
8	652	79.4	152	1 DSHOCZ	superoxide dismuta
9	650.5	79.2	152	1 DSHOCZ	superoxide dismuta
10	635	77.3	151	2 A24475	superoxide dismuta
11	599	73.0	153	2 S65436	superoxide dismuta
12	589.5	71.8	166	2 S29782	superoxide dismuta
13	558.5	68.0	152	2 S04623	superoxide dismuta
14	536.5	65.3	151	1 DSHOCZ	superoxide dismuta
15	533	64.9	151	1 S09568	superoxide dismuta
16	521	63.5	150	2 S05021	superoxide dismuta
17	512	62.4	153	2 A45171	superoxide dismuta
18	505	61.5	153	2 S48117	superoxide dismuta
19	493	60.0	153	1 DSHOCZ	superoxide dismuta
20	492	60.0	153	2 S05498	superoxide dismuta
21	492	59.9	153	2 A49241	superoxide dismuta
22	487	59.3	153	2 S03606	superoxide dismuta
23	482.5	58.8	154	2 S20511	superoxide dismuta
24	480	58.5	151	2 A29077	superoxide dismuta
25	477.5	58.2	152	2 S21136	superoxide dismuta
26	469.5	57.2	152	2 S07007	superoxide dismuta
27	465.5	56.7	201	2 T06800	superoxide dismuta
28	464.5	56.6	216	2 T51730	superoxide dismuta
29	464.5	56.6	222	1 DSSPCZ	superoxide dismuta

30	463.5	56.5	202	1 DSPMCZ	superoxide dismuta
31	463.5	56.5	211	2 T03685	probable superoxid
32	462.5	56.3	152	2 S22508	superoxide dismuta
33	462.5	56.3	201	2 T06229	probable superoxid
34	462	56.3	158	2 S18743	superoxide dismuta
35	461.5	56.2	152	2 S72235	superoxide dismuta
36	460	56.0	158	2 A48256	superoxide dismuta
37	455.5	55.5	156	2 T12204	superoxide dismuta
38	455.5	55.5	219	2 S03608	superoxide dismuta
39	453.5	55.2	217	2 S48021	superoxide dismuta
40	451.5	55.0	152	2 T06570	superoxide dismuta
41	451.5	55.0	162	2 T51731	superoxide dismuta
42	450.5	54.9	218	2 H84681	probable copper/z1
43	443.5	54.0	152	2 S40404	superoxide dismuta
44	441.5	53.8	152	2 JW0084	superoxide dismuta
45	439.5	53.5	152	1 DSSPCZ	superoxide dismuta

ALIGNMENTS

RESULT 1

DSHUCZ

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Aug-1980 #sequence_revision 19-Feb-1984 #text_change 21-Jul-2000

C:Accession: A22703; PC2217; A23046; A93969; JX0055; A91282; A90440; T54382; I68676;

R:Levanon, D.; Lieman-Hurwitz, J.; Dafni, N.; Wigderson, M.; Sherman, L.; Bernstein,

EMBO J. 4, 77-84, 1985

A:Title: Architecture and anatomy of the chromosomal locus in human chromosome 21 enc

A:Reference number: A22703; MUID:85257452; PMID:3160582

A:Accession: A22703

A:Molecule type: DNA

A:Residues: 2-154 <LEV>

A:Cross-references: EMBL:X01780; NID:g36534; PIDN:CAA25915.1; PID:g36535; EMBL:X01781

A:Note: the authors did not translate the codon for residue 1

R:Kim, H.T.; Kim, Y.H.; Nam, J.W.; Lee, H.J.; Rho, H.M.; Jung, G.

Biochem. Biophys. Res. Commun. 201, 1526-1533, 1994

A:Title: Study of 5'-flanking region of human Cu/Zn superoxide dismutase.

A:Reference number: PC2217; MUID:94296434; PMID:8024598

A:Accession: PC2217

A:Molecule type: DNA

A:Residues: 1-24 <KIM>

A:Cross-references: EMBL:Z29336; NID:g531099

R:Hallewell, R.A.; Maslars, F.R.; Najarian, R.C.; Puma, J.P.; Quiroga, M.R.; Randolph

Nucleic Acids Res. 13, 2017-2034, 1985

A:Title: Human Cu/Zn superoxide dismutase cDNA: isolation of clones synthesising high

A:Reference number: A23046; MUID:85215596; PMID:3889846

A:Accession: A23046

A:Molecule type: mRNA

A:Residues: 1-154 <HAL>

A:Cross-references: EMBL:X02317; NID:g36541; PIDN:CAA26182.1; PID:g36542

R:Sherman, L.; Dafni, N.; Lieman-Hurwitz, J.; Groner, Y.

Proc. Natl. Acad. Sci. U.S.A. 80, 5465-5469, 1983

A:Title: Nucleotide sequence and expression of human chromosome 21-encoded superoxide

A:Reference number: A93969; MUID:83299994; PMID:6577438

A:Accession: A93969

A:Molecule type: mRNA

A:Residues: 2-154 <SHE>

A:Cross-references: GB:X00065; NID:g36541; PIDN:CAA26182.1; PID:g36542

A:Note: the authors did not translate the codon for residue 1

R:Kajihara, J.; Enomoto, M.; Nishijima, K.; Yabuuchi, M.; Katoh, K.

J. Biochem. 104, 851-854, 1988

A:Title: Comparison of properties between human recombinant and placental copper-zinc

A:Reference number: JX0055; MUID:89174523; PMID:2853161

A:Accession: JX0055

A:Molecule type: protein

A:Residues: 2-154 <KAJ>

A:Experimental source: placenta

R:Barra, D.; Martini, F.; Bannister, J.V.; Schinina, M.E.; Rotilio, G.; Bannister, W.

FEBS Lett. 120, 53-56, 1980

A:Title: The complete amino acid sequence of human Cu/Zn superoxide dismutase.

A:Reference number: A91282; MUID:81067132; PMID:7002610

A:Accession: A91282
 A:Molecule type: protein
 A:Residues: 2-11,'N',13-92,'N',94-154 <BAR>
 A:Experimental source: erythrocytes
 R:Jabusch, J.R.; Farb, D.L.; Kerschensteiner, D.A.; Deutsch, H.F.
 Biochemistry 19, 2310-2316, 1980
 A:Title: Some sulphydryl properties and primary structure of human erythrocyte superoxide
 A:Reference number: A90440; MUID:80221052; PMID:6770891
 A:Accession: A90440
 A:Molecule type: protein
 A:Residues: 2-17,'S',19-26,'N',28-49,'Q',51-52,'ND',55-98,'V',100-154 <TAB>
 R:Enayati, Z.E.; Orrell, R.W.; Claus, A.; Ludolph, A.; Bachus, R.; Brockmueller, J.; Ray-
 cell, J.
 Hum. Mol. Genet. 4, 1239-1240, 1995
 A:Title: Two novel mutations in the gene for copper zinc superoxide dismutase in UK fami
 A:Reference number: I54382; MUID:96133303; PMID:8528216
 A:Accession: I54382
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'SMKQIH',127-154 <ENAL>
 A:Cross-references: GB:I44746; NID:g928824; PIDN:AAC41773.1; PID:g928825
 A:Note: mutated exon 5 from patient with familial amyotrophic lateral sclerosis
 A:Accession: I68676
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 25-48,'Q',50-56 <ENAZ>
 A:Cross-references: GB:I46374; NID:g939772; PIDN:AAB59626.1; PID:g939773
 A:Note: mutant sequence from patient with familial amyotrophic lateral sclerosis
 A:Accession: I68677
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 120-125,'H',127-154 <ENAZ>
 A:Cross-references: GB:I46375; NID:g939774; PIDN:AAB59627.1; PID:g939775
 A:Note: mutant sequence from patient with familial amyotrophic lateral sclerosis
 R:Yamazaki, Y.; Takao, T.; Murata, H.; Kawaharada, Y.; Sugiyama, T.; Shimonishi, Y.
 Res. Commun. Biochem. Cell Mol. Biol. 1, 205-217, 1997
 A:Title: The use of capillary liquid chromatography/electrospray mass spectrometry to id
 A:Reference number: JC5686
 A:Accession: JC5686
 A:Molecule type: protein
 A:Residues: 2-154 <YAM>
 A:Note: Inter-molecular disulfide bonds form under non-physiological conditions
 R:Rosen, D.R.; Siddique, T.; Patterson, D.; Figlewicz, D.A.; Sapp, P.; Hentati, A.; Dona
 n, S.M.; Berger, R.; Tanzi, R.E.; Halperin, J.J.; Herzfeldt, B.; Van den Bergh, R.; Hung
 Rouleau, G.A.; Gusella, J.S.; Horvitz, H.R.; Brown Jr., R.H.
 Nature 362, 59-62, 1993
 A:Title: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyot
 A:Reference number: A44737; MUID:93188938; PMID:8446170
 A:Contents: annotation; mutant sequences reported
 R:Pargie, H.E.; Tainer, J.A.
 submitted to the Brookhaven Protein Data Bank, July 1993
 A:Reference number: A52244; PDB:1SPD
 A:Contents: annotation; X-ray crystallography, 2.4 angstroms, residues 2-154
 A:Note: erythrocyte recombinant form expressed in yeast
 R:Pargie, H.E.; Hallewell, R.A.; Tainer, J.A.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6109-6113, 1992
 A:Title: Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn
 A:Reference number: A44540; MUID:92335247; PMID:1463506
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms
 C:Genetics:
 A:Gene: GDB:SOD1; ALS; ALS1
 A:Cross-references: GDB:119596; OMIM:147450
 A:Map position: 21q22.1-21q22.1
 A:Introns: 24/3; 56/3; 80/3; 119/3
 A:Note: a defect in this enzyme is the cause of adult familial amyotrophic lateral scler
 C:Function:
 A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
 C:Superfamily: superoxide dismutase (Cu-Zn)
 C:Keywords: acetylated amino end; copper; homodimer; Lou Gehrig's disease; metalloprotei
 F:2-154/Product: superoxide dismutase (Cu-Zn) #status experimental <MAT>
 F:2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental

F:47,49,64,121/Binding site: copper (His) #status experimental
 F:58-147/Disulfide bonds: #status experimental
 F:64,72,81,84/Binding site: zinc (His, His, Asp) #status experimental
 F:144/Active site: Arg #status experimental
 Query Match 100.0%; Score 821; DB 1; Length 154;
 Best Local Similarity 100.0%; Pred. No. 1.6e-69;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
 DB 1 MATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
 QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 RESULT 2
 T43640
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - synthetic (fragment)
 C:Species: synthetic
 A:Note: human SOD1 and SOD3 genes engineered and expressed in Schizosaccharomyces pom
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 08-Dec-2000
 C:Accession: T43640
 R:Inoue, M.; Watanabe, N.; Matsuno, K.; Sasaki, J.; Tanaka, Y.; Hatanaka-Sawai, H.; A
 J. Biol. Chem. 266, 16409-16414, 1991
 A:Title: Expression of a hybrid Cu/Zn-type superoxide dismutase which has high affinity
 A:Reference number: Z22590; MUID:91358420; PMID:1885572
 A:Accession: T43640
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-179 <INO>
 A:Cross-references: EMBL:D90358; NID:g218563; PIDN:BAAL4373.1; PID:g218564
 A:Experimental source: strain EH13-15
 A:Note: the engineered initiator codon and the termination codon are not shown
 C:Comment: For the original sequences, see PIR:DSHUCZ and PIR:DSHUEC.
 C:Keywords: copper; metalloprotein; oxidoreductase; zinc
 F:1-151/Region: SOD1 gene derived
 F:153-179/Region: SOD3 gene derived
 F:46,48,63,120/Binding site: copper (His) #status predicted
 F:57-146/Disulfide bonds: #status predicted
 F:63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted
 F:143/Active site: Arg #status predicted

Query Match 98.3%; Score 807; DB 4; Length 179;
 Best Local Similarity 100.0%; Pred. No. 3.8e-68;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
 DB 1 ATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
 QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 121
 DB 61 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
 QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGI 152
 DB 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGI 151
 RESULT 3
 S36108
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 19-Mar-1997 #sequence_revision 09-Aug-1997 #text_change 20-Apr-2000
 C:Accession: S36108; S55796
 R:Wolf, B.; Reinecke, K.; Ammann, K.D.; Brigelius-Flohe, R.; Flohe, L.

[illegible]

RESULT 8

DSBOCZ
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - bovine
N:Alternate names: thymus hypocholesterolemic factor
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 27-Feb-1997 #text_change 16-Jun-2000
C:Accession: I45883; S13674; A92163; A56902; A00513
R:Halliwell, R.A.; Imlay, K.C.; Lee, P.; Fong, N.M.; Gallegos, C.; Getzoff, E.D.; Tainer
Biochem. Biophys. Res. Commun. 181, 474-480, 1991
A:Title: Thermostabilization of recombinant human and bovine CuZn superoxide dismutases
A:Reference number: I45883; MUID:92068236; PMID:1958215
A:Accession: I45883
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-152 <HAL>
A:Cross-references: GB:M81129; NID:g162960; PIDN:AAA73164.1; PID:g162961
R:Gibbs, L.S.; Shaffer, J.B.
Nucleic Acids Res. 18, 7171, 1990
A:Title: Nucleotide sequence of bovine copper/zinc superoxide dismutase cDNA generated by
A:Reference number: S13674; MUID:91088332; PMID:2263495
A:Accession: S13674
A:Molecule type: mRNA
A:Residues: 2-152 <GIB>
A:Cross-references: EMBL:X54799
A:Experimental source: lung
R:Steinman, H.M.; Naik, V.R.; Abernethy, J.L.; Hill, R.L.
J. Biol. Chem. 249, 7326-7338, 1974
A:Title: Bovine erythrocyte superoxide dismutase. Complete amino acid sequence.
A:Reference number: A92163; MUID:75060425; PMID:4279916
A:Accession: A92163
A:Molecule type: protein
A:Residues: 2-152 <STE>
R:Abernethy, J.L.; Steinman, H.M.; Hill, R.L.
J. Biol. Chem. 249, 7339-7347, 1974
A:Title: Bovine erythrocyte superoxide dismutase. Subunit structure and sequence location
A:Reference number: A92164; MUID:75060426; PMID:4436313
A:Contents: annotation; disulfide bond
A>Note: the molecule is a dimer of identical chains
R:Richardson, J.S.; Thomas, K.A.; Rubin, B.H.; Richardson, D.C.
Proc. Natl. Acad. Sci. U.S.A. 72, 1349-1353, 1975
A:Title: Crystal structure of bovine Cu,Zn superoxide dismutase at 3 angstrom resolution
A:Reference number: A93796; MUID:75158289; PMID:1055410
A:Contents: annotation; X-ray crystallography, 3.0 angstroms
R:Tainer, J.A.; Getzoff, E.D.; Richardson, J.S.; Richardson, D.C.
submitted to the Brookhaven Protein Data Bank, March 1980
A:Reference number: A50536; PDB:2SOD
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 1-151
R:Malinowski, D.P.; Friedovich, I.
Biochemistry 18, 5909-5917, 1979
A:Title: Chemical modification of arginine at the active site of the bovine erythrocyte
A:Reference number: A90436; MUID:80088307; PMID:518876
A:Contents: annotation; active site
A>Note: chemical modification of Arg-142 inactivates the enzyme
R:Mondola, P.; Santillo, M.; Belfiore, A.; Camardella, L.; Santangelo, F.
Comp. Biochem. Physiol. B 104, 193-199, 1993
A:Title: The thymus hypocholesterolemic factor (TphF): a bovine thymic superoxide dismut
A:Reference number: A56902; MUID:93193420; PMID:8448990
A:Accession: A56902
A:Molecule type: protein
A:Residues: 5-38;70-130;136-152 <MON>
A>Note: sequence modified after extraction from NCBI backbone
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Cu-Zn)
C:Keywords: acetylated amino end; copper; homodimer; metalloprotein; oxidoreductase; zinc
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status experimental
F:45,47,62,119/Binding site: copper (His) #status experimental
F:56-145/disulfide bonds: #status experimental
F:62,70,79,82/Binding site: zinc (His, His, Asp) #status experimental
F:142/Active site: Arg #status experimental

Query Match 79.4%; Score 652; DB 1; Length 152;

Best Local Similarity 81.8%; Pred. No. 9.4e-54;

Matches 126; Conservative 10; Mismatches 16; Indels 2; Gaps 1;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60

DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 58

QY 61 AGPHNPISLRKHGPKDERHVGDLGNVTADKGDVADVSIEDSVLSGDHCHIGRTLVW 120

DB 59 AGPHNPISLRKHGPKDERHVGDLGNVTADKNGVAVIYDIDPLTSLSGEYSIIIGRTMVV 118

QY 121 HEKADDLGKGGNESTKTGNAGSRLACGVIGTAA 154

DB 119 HEKPDLLGRGGNESTKTGNAGSRLACGVIGTAA 152

RESULT 9

DSBOCZ

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) [validated] - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 20-Apr-2000

C:Accession: A00514

R:Schinina, M.E.; Barra, D.; Simmaco, M.; Bossa, F.; Rottilio, G.

FEBS Lett. 186, 267-270, 1985

A:Title: Primary structure of porcine Cu,Zn superoxide dismutase.

A:Reference number: A00514; MUID:85231234; PMID:3891411

A:Accession: A00514

A:Molecule type: protein

A:Residues: 1-152 <SCH>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg

C:Superfamily: superoxide dismutase (Cu-Zn)

C:Keywords: acetylated amino end; copper; metalloprotein; oxidoreductase; zinc

F:1/Modified site: acetylated amino end (Ala) #status experimental

F:45,47,62,119/Binding site: copper (His) #status predicted

F:56-145/disulfide bonds: #status predicted

F:62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted

F:142/Active site: Arg #status predicted

Query Match 79.2%; Score 650.5; DB 1; Length 152;

Best Local Similarity 82.4%; Pred. No. 1.3e-53;

Matches 126; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSA 61

DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSA 59

QY 62 GFHNPISLRKHGPKDERHVGDLGNVTADKGDVADVSIEDSVLSGDHCHIGRTLVVH 121

DB 60 GFHNPISLRKHGPKDERHVGDLGNVTADKGDVADVSIEDSVLSGDHCHIGRTLVVH 119

QY 122 EKADDLGKGGNESTKTGNAGSRLACGVIGTAA 154

DB 120 EKPDLLGRGGNESTKTGNAGSRLACGVIGTAA 152

RESULT 10

A24475

superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 05-Mar-1999

C:Accession: A24475

R:Schinina, M.E.; Barra, D.; Gentilomo, S.; Bossa, F.; Capo, C.; Rottilio, G.; Calabre

FEBS Lett. 207, 7-10, 1986

A:Title: Primary structure of a cationic Cu,Zn superoxide dismutase. The sheep enzyme

A:Reference number: A24475; MUID:87030907; PMID:3770194

A:Accession: A24475

A:Molecule type: protein

A:Residues: 1-151 <SCH>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg

C:Superfamily: superoxide dismutase (Cu-Zn)

C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F:44,46,61,118/Binding site: copper (His) #status predicted
F:55-144/Disulfide bonds: #status predicted
F:61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted
F:141/Active site: Arg #status predicted

Query Match 77.3%; Score 635; DB 2; Length 151;
Best Local Similarity 80.9%; Pred. No. 3.6e-52;
Matches 123; Conservative 8; Mismatches 19; Indels 2; Gaps 1;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDGPVQGTIRFEAK--GDKVVVVTGSLTEGDHGFHVFHFGDNTAGCTSA 58
QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSTEDSVISLSDHCHCIIGRTLVVH 121
DB 59 GPHFNPLSRKHGPKDEERHVGDLGNVADKNGVAIVDVLPLSLSGEYSIIIGRTMVVH 118
QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIA 153
DB 119 EKPDLLGRGGNEESTKTGNAGSLACGVIGIA 150

RESULT 11
S65436
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - chicken
C;Species: Gallus gallus (chicken)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Apr-2000
C;Accession: S65436
R;Schinina, M.E.; Carlini, P.; Polticelli, F.; Zappacosta, F.; Bossa, F.; Calabrese, L.
Eur. J. Biochem. 237, 433-439, 1996
A;Title: Amino acid sequence of chicken Cu, Zn-containing superoxide dismutase and ident
A;Reference number: S65436; MUID:96215439; PMID:8647082
A;Accession: S65436
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-153 <SCH>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F:46,48,63,119/Binding site: copper (His) #status predicted
F:57-145/Disulfide bonds: #status predicted
F:63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted
F:142/Active site: Arg #status predicted

Query Match 73.0%; Score 599; DB 2; Length 153;
Best Local Similarity 74.2%; Pred. No. 8.4e-49;
Matches 112; Conservative 19; Mismatches 18; Indels 2; Gaps 2;

QY 4 KAVCVLKGDGPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTSAGP 63
DB 4 KAVCVLMKGDGPVGVHIFQO--QSGSPKVTGKITGLSDGDHGFHVFHFGDNTAGCTSAGA 62
QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSTEDSVISLSDHCHCIIGRTLVVH 123
DB 63 HFNPEKGQHGPKDADRHVVDLGNVTA--KGGVAEVEIEDSVISLTPCHCIIGRTMVVHAK 121
QY 124 ADDLKGKGGNEESTKTGNAGSLACGVIGIA 154
DB 122 SDDLGRGGNESKLTGNAGPRLACGVIGIA 152

RESULT 12
S29782
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - loggerhead
C;Species: Caretta caretta (loggerhead)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-May-2000
C;Accession: S29782
R;Schinina, M.E.; Bossa, F.; Lania, A.; Capo, C.R.; Carlini, P.; Calabrese, L.
Eur. J. Biochem. 211, 843-849, 1993
A;Title: The primary structure of turtle Cu,Zn superoxide dismutase. Structural and func
A;Reference number: S29782; MUID:93170321; PMID:8436140

A;Accession: S29782
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-166 <SCH>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F:58,60,75,132/Binding site: copper (His) #status predicted
F:69-158/Disulfide bonds: #status predicted
F:75,83,92,95/Binding site: zinc (His, His, Asp) #status predicted
F:155/Active site: Arg #status predicted

Query Match 71.8%; Score 589.5; DB 2; Length 166;
Best Local Similarity 69.9%; Pred. No. 7.2e-48;
Matches 114; Conservative 16; Mismatches 20; Indels 13; Gaps 2;

QY 4 KAVCVLKGDG-----PVGIIINFEQKESNGPKVWGSIKGLTEGLHGFHVFH 51
DB 4 KAVCVLKGEDPVKEPVKGPVKIYFEQ--QNGPVTLSGSLITGLTEGKHGFHVFH 62
QY 52 GDNATAGCTSAGPHNPLSRKHGPKDEERHVGDLGNVTADKDGADVSTEDSVISLSDH 111
DB 63 GDNATAGCTSAGAHFNPPGKNHGGPDNERHVGDLGNVIANKEGVAEVCIKDSLISLTGSQ 122
QY 112 CIIGRTLIVVHEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
DB 123 SIIGRTWVHEKEDDLGKGGNDESLKTGNAGSLACGVIGIAK 165

RESULT 13
S04623
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - blue shark
C;Species: Prionace glauca (blue shark)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Apr-2000
C;Accession: S04623; S45642
R;Calabrese, L.; Polticelli, F.; O'Neill, P.; Galtierl, A.; Barra, D.; Schinina, E.;
FEBS Lett. 250, 49-52, 1989
A;Title: Substitution of arginine for lysine 134 alters electrostatic parameters of t
A;Reference number: S04623; MUID:89250032; PMID:2500367
A;Accession: S04623
A;Molecule type: protein
A;Residues: 1-152 <CAL>
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C;Superfamily: superoxide dismutase (Cu-Zn)
C;Keywords: copper; metalloprotein; oxidoreductase; zinc
F:44,46,61,118/Binding site: copper (His) #status predicted
F:55-144/Disulfide bonds: #status predicted
F:61,69,78,81/Binding site: zinc (His, His, Asp) #status predicted
F:141/Active site: Arg #status predicted

Query Match 68.0%; Score 558.5; DB 2; Length 152;
Best Local Similarity 70.9%; Pred. No. 5.1e-45;
Matches 107; Conservative 15; Mismatches 28; Indels 1; Gaps 1;

QY 4 KAVCVLKGDGPVQGIINFEQKESNGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTSAGP 63
DB 2 KAVCVLKGTGVTGTVLFEQ--ADGPVTLKGSITGLTPGKHGFHVFHFGDNTAGCTSAGP 60
QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKDGADVSTEDSVISLSDHCHCIIGRTLVVH 123
DB 61 HYNPFSKNHGGPDDEERHVGDLGNVEANGVAFKIDRQLHLSGRSIIIGRTLVVH 120
QY 124 ADDLKGKGGNEESTKTGNAGSLACGVIGIAQ 154
DB 121 EDDLKGKGGDEESLRTGNAGSLACGVIGIAK 151

RESULT 14
DSWFCZ
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) - swordfish
C;Species: Xiphias gladius (swordfish)

C>Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Mar-1999

C:Accession: A00516

R:Rocha, H.A.; Bannister, W.H.; Bannister, J.V.

Eur. J. Biochem. 145, 477-484, 1984

A:Title: The amino-acid sequence of copper/zinc superoxide dismutase from swordfish liver

A:Reference number: A00516; PMID:85076642; PMID:6510412

A:Accession: A00516

A:Molecule type: protein

A:Residues: 1-151 <ROC>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C:Superfamily: superoxide dismutase (Cu-Zn)

C:Keywords: copper; metalloprotein; oxidoreductase; zinc

F:46,48,63,119/Binding site: copper (His) #status predicted

F:57-145/Disulfide bonds: #status predicted

F:63,71,80,83/Binding site: zinc (His, His, Asp) #status predicted

F:142/Active site: Arg #status predicted

Query Match 65.3%; Score 536.5; DB 1; Length 151;

Best Local Similarity 69.6%; Pred. No. 5.7e-43;

Matches 103; Conservative 14; Mismatches 30; Indels 1; Gaps 1;

QY 4 KAVCVLKGDPVQGIINFEOKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSAGP 63

DB 3 KAVCVLAGSGDVKGVVHFEQQD-EGAVSVEGRKEGLTDGLHGFHIVHFGDNTGCMGSGS 61

QY 64 HFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHCIIGRTLVVHEK 123

DB 63 HFNPAKKHAGPKDEDRHVGDLGNVTADANGVAKIDITDK-ISLTGPGYIIIGRTMVIHEK 121

QY 124 ADDLKGKGNBESLTKGNAGSLACGVIG 151

DB 122 ADDLGRGGBESLTKGNAGSLACGVIG 149

RESULT 15

S09568

C:Species: Xenopus laevis (African clawed frog)

C>Date: 13-Jan-1995 #sequence_revision 11-Apr-1997 #text_change 11-Jun-1999

C:Accession: S09568; S05022; S59617

R:Carri, M.T.; Battistoni, A.; Mariottini, P.; Rotilio, G.

Nucleic Acids Res. 18, 1641, 1990

A:Title: Xenopus laevis Cu,Zn superoxide dismutase B cDNA sequence.

A:Reference number: S09568; PMID:90221905; PMID:2326205

A:Accession: S09568

A:Molecule type: mRNA

A:Residues: 1-151 <CAR>

A:Cross-references: EMBL:X51518; NID:g65258; PIDN:CAA35890.1; PID:g65259

R:Schinina, M.E.; Barra, D.; Bossa, F.; Calabrese, L.; Montesano, L.; Carri, M.T.; Mario

Arch. Biochem. Biophys. 272, 507-515, 1989

A:Title: Primary structure from amino acid and cDNA sequences of two Cu,Zn superoxide di

A:Reference number: S05021; PMID:89321563; PMID:2751312

A:Accession: S05022

A:Molecule type: protein

A:Residues: 2-60, 'P', 62-151 <SCH>

A>Note: residues 78-151 were also determined by cDNA sequencing

A:Accession: S59617

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 78-151 <SC2>

C:Function:

A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C:Superfamily: superoxide dismutase (Cu-Zn)

C:Keywords: copper; metalloprotein; oxidoreductase; zinc

F:45,47,62,118/Binding site: copper (His) #status predicted

F:56-144/Disulfide bonds: #status predicted

F:62,70,79,82/Binding site: zinc (His, His, Asp) #status predicted

F:141/Active site: Arg #status predicted

Query Match

Best Local Similarity 64.9%; Score 533; DB 1; Length 151;

Matches 102; Conservative 16; Mismatches 28; Indels 2; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEOKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTSAGP 63
DB 3 KAVCVLAGSGDVKGVVHFEQQD-EGAVSVEGRKEGLTDGLHGFHIVHFGDNTGCMGSGS 61
QY 64 HFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHCIIGRTLVVHEK 123
DB 62 HFNPNKNHAGAPGDDTRHVGDLGNVTAE-GGVAQFKITDSLISLKGPNISIGRTAVVHEK 120
QY 124 ADDLKGKGNBESLTKGNAGSLACGVIG 151
DB 121 ADDLKGKGNBESLTKGNAGSLACGVIG 148

Search completed: January 28, 2003, 14:20:13
Job time : 5.03824 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:03:18 ; Search time 2.35564 Seconds
(without alignments)
2711.515 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATRAVCVLKGDGPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	816	99.4	153	1 SODC_HUMAN	P00441 homo sapien
2	684.5	83.4	152	1 SODC_CAVPO	P33431 cavia porce
3	675	82.2	153	1 SODC_MOUSE	P08228 mus musculu
4	673	82.0	153	1 SODC_RAT	P07632 rattus norv
5	662	80.6	153	1 SODC_HORSE	P00443 equus cabal
6	654.5	79.7	152	1 SODC_RABIT	P09212 oryctolagus
7	651	79.3	151	1 SODC_CEREL	O46412 cervus elap
8	650.5	79.2	152	1 SODC_PIG	P04178 sus scrofa
9	647	78.8	151	1 SODC_BOVIN	P00442 bos taurus
10	635	77.3	151	1 SODC_SHEEP	P09670 ovis aries
11	593	72.2	153	1 SODC_CHICK	P00566 gallus gall
12	589.5	71.8	166	1 SODC_CARCR	P00174 caretta car
13	579	70.5	154	1 SODC_BRARE	O73872 brachydanio
14	558.5	68.0	152	1 SODC_PRIGL	P1418 prionace gl
15	541	65.9	150	1 SODD_XENLA	P15107 xenopus lae
16	536.5	65.3	151	1 SODC_XIPGL	P03946 xiphias gla
17	521	63.5	150	1 SODC_XENLA	P13926 xenopus lae
18	507	61.8	152	1 SODC_CERCA	P28755 ceratitlis c
19	500	60.9	152	1 SODC_CHYAM	Q07182 chymomyza a
20	497	60.5	153	1 SODC_SCHMA	Q11137 schistosoma
21	495	60.3	152	1 SODC_DROWI	P41973 drosophila
22	493.5	60.1	159	1 SODC_HAECO	Q27666 haemochus
23	488	59.4	152	1 SODC_DROME	P00444 drosophila
24	483.5	58.9	153	1 SODC_DEBHA	O42724 debaryomyce
25	482.5	58.8	154	1 SODC_PINSY	P24669 pinus sylve
26	482	58.7	152	1 SODC_DROVI	P10791 drosophila
27	480	58.5	150	1 SODC_MAIZE	P11428 zea mays (m
28	477.5	58.2	151	1 SOD2_ORYSA	P28757 oryza sativ
29	474	57.7	158	1 SODC_BRUPA	P41962 brugia paha
30	473.5	57.7	220	1 SODP_SOLCS	O04997 solidago ca
31	472.5	57.6	151	1 SODC_HALRO	P81926 halocynthia
32	472.5	57.6	153	1 SODC_ASPTU	Q9y849 aspergillus
33	470.5	57.3	151	1 SOD5_MAIZE	P23346 zea mays (m

RESULT 1

ID	SODC_HUMAN	STANDARD;	PRT;	153 AA.
AC	P00441; Q16711; Q16669; Q16838; Q16840;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).			
GN	SOD1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85257452; PubMed=3160582;			
RA	Levanon D., Lieman-Hurwitz J., Dafni N., Wigderson M., Sherman L.,			
RA	Bernstein Y., Laver-Rudich Z., Danciger E., Stein O., Groner Y.;			
RT	"Architecture and anatomy of the chromosomal locus in human			
RT	chromosome 21 encoding the Cu/Zn superoxide dismutase.";			
RL	EMBO J. 4:77-84(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215596; PubMed=3889846;			
RA	Hallewell R.A., Masiazk F.R., Najarian R.C., Puma J.P., Quiroga M.R.,			
RA	Randolph A., Sanchez-Pescador R., Scandella C.J., Smith B.,			
RA	Steimer K.S., Mullenbach G.T.;			
RT	"Human Cu/Zn superoxide dismutase cDNA: isolation of clones			
RT	synthesising high levels of active or inactive enzyme from an			
RT	expression library.";			
RL	Nucleic Acids Res. 13:2017-2034(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83299994; PubMed=6577438;			
RA	Sherman L., Dafni N., Lieman-Hurwitz J., Groner Y.;			
RT	"Nucleotide sequence and expression of human chromosome 21-encoded			
RT	superoxide dismutase mRNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 80:5465-5469(1983).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89174523; PubMed=2853161;			
RA	Kajihara J., Enomoto M., Nishijima K., Yabuuchi M., Katoh K.;			
RT	"Comparison of properties between human recombinant and placental			
RT	copper-zinc SOD.";			
RL	J. Biochem. 104:851-854(1988).			
RN	[5]			
RP	SEQUENCE.			
RX	MEDLINE=810671132; PubMed=7002610;			
RA	Barra D., Martini F., Bannister J.V., Schinina M.E., Rotillo G.,			
RA	Bannister W.H., Bossa F.;			
RT	"The complete amino acid sequence of human Cu/Zn superoxide			
RT	dismutase.";			
RL	FEBS Lett. 120:53-56(1980).			
RN	[6]			
RP	SEQUENCE.			
RX	MEDLINE=80221052; PubMed=6770891;			

P23345 zea mays (m
O65198 medicago sa
P54407 drosophila
O22668 panax glnse
O07449 onchocerca
O78310 arabidopsis
P07505 spinacia ol
P11984 pisum sativ
P93407 oryza sativ
P28756 oryza sativ
P24706 onchocerca
Q98q15 ananas como

ALIGNMENTS

RA Jabusch J.R., Farb D.L., Kerschensteiner D.A., Deutsch H.F.;
 RT "some sulfhydryl properties and primary structure of human
 RT erythrocyte superoxide dismutase.";
 RL Biochemistry 19:2310-2316(1980).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=92335247; PubMed=1463506;
 RA Parge H.E., Hallewell R.A., Tainer J.A.;
 RA "Atomic structures of wild-type and thermostable mutant recombinant
 RT human Cu,Zn superoxide dismutase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6109-6113(1992).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF VARIANT GLY-37.
 RX MEDLINE=98200458; PubMed=9541385;
 RA Hart P.J., Liu H., Pellegrini M., Nersissian A.M., Gralla E.B.,
 RA Valentine J.S., Eisenberg D.;
 RA "Subunit asymmetry in the three-dimensional structure of a human
 RT Cu,ZnSOD mutant found in familial amyotrophic lateral sclerosis.";
 RL Protein Sci. 7:545-555(1998).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (1.02 ANGSTROMS).
 RX MEDLINE=99262851; PubMed=10329151;
 RA Ferraroni M., Rypniewski W., Wilson K.S., Viezzoli M.S., Banci L.,
 RA Bertini I., Mangani S.;
 RA "The crystal structure of the monomeric human SOD mutant
 RT F50E/G51E/E133Q at atomic resolution. The enzyme mechanism
 RT revisited.";
 RL J. Mol. Biol. 288:413-426(1999).
 RN [10]
 RP STRUCTURE BY NMR.
 RX MEDLINE=98384143; PubMed=9718300;
 RA Banci L., Benedetto M., Bertini I., del Conte R., Piccioli M.,
 RA Viezzoli M.S.;
 RA "Solution structure of reduced monomeric Q133M2 copper, zinc
 RT superoxide dismutase (SOD). Why is SOD a dimeric enzyme?;
 RL Biochemistry 37:11780-11791(1998).
 RN [11]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=96154858; PubMed=8592323;
 RA de Belleroche J., Orrell R., King A.;
 RA "Familial amyotrophic lateral sclerosis/motor neurone disease (FALS):
 RT a review of current developments.";
 RL J. Med. Genet. 32:841-847(1995).
 RN [12]
 RP VARIANTS FALS.
 RX MEDLINE=93188958; PubMed=8446170;
 RA Rosen D.R., Siddique T., Patterson D., Figlewicz D.A., Sapp P.,
 RA Rentati A., Donaldson D., Goto J., O'Regan J.P., Deng H.-X.,
 RA Rahmani Z., Krizus A., McKenna-Yasek D., Cayabyab A., Gaston S.M.,
 RA Berger R., Tanzi R.E., Halperin J.J., Herzfeldt B., van den Bergh R.,
 RA Hung W.-Y., Bird T.E., Deng G., Mulder D.W., Smyth C., Laing N.G.,
 RA Soriano E., Pericak-Vance M.A., Haines J., Rouleau G.A., Gusella J.S.,
 RA Horvitz H.R., Brown R.H. Jr.;
 RA "Mutations in Cu/Zn superoxide dismutase gene are associated with
 RT familial amyotrophic lateral sclerosis.";
 RL Nature 362:59-62(1993).
 RN [13]
 RP ERRATUM.
 RX MEDLINE=93323981; PubMed=8332197;
 RA Rosen D.R.;
 RA Nature 364:362-362(1993).
 RN [14]
 RP VARIANTS FALS.
 RX MEDLINE=93355289; PubMed=8351519;
 RA Deng H.-X., Rentati A., Tainer J.A., Iqbal Z., Cayabyab A.,
 RA Hung W.-Y., Getzoff E.D., Hu P., Herzfeldt B., Roos R.P., Warner C.,
 RA Deng G., Soriano E., Smyth C., Parge H.E., Ahmed A., Roses A.D.,
 RA Hallewell R.A., Pericak-Vance M.A., Siddique T.;
 RA "Amyotrophic lateral sclerosis and structural defects in Cu,Zn
 RT superoxide dismutase.";
 RL Science 261:1047-1051(1993).
 RN [15]
 RP VARIANT FALS THR-4.

RX MEDLINE=94235014; PubMed=8179602;
 RA Nakano R., Sato S., Inuzuka T., Sakimura K., Mishina M., Takahashi H.,
 RA Ikuta F., Honma Y., Fujii J., Taniguchi N., Tsuji S.;
 RA "A novel mutation in Cu/Zn superoxide dismutase gene in Japanese
 RT familial amyotrophic lateral sclerosis.";
 RL Biochem. Biophys. Res. Commun. 200:695-703(1994).
 RN [16]
 RP VARIANT FALS GLU-7.
 RX MEDLINE=95071364; PubMed=7980516;
 RA Hirano M., Fujii J., Nagai Y., Sonobe M., Okamoto K., Araki H.,
 RA Taniguchi N., Ueno S.;
 RA "A new variant Cu/Zn superoxide dismutase (Val7-->Glu) deduced from
 RT lymphocyte mRNA sequences from Japanese patients with familial
 RT amyotrophic lateral sclerosis.";
 RL Biochem. Biophys. Res. Commun. 204:572-577(1994).
 RN [17]
 RP VARIANT FALS LYS-21.
 RX MEDLINE=94348517; PubMed=8069312;
 RA Jones C.T., Swinger R.J., Brock D.J.H.;
 RA "Identification of a novel SOD1 mutation in an apparently sporadic
 RT amyotrophic lateral sclerosis patient and the detection of Ile113Thr
 RT in three others.";
 RL Hum. Mol. Genet. 3:649-650(1994).
 RN [18]
 RP VARIANTS FALS ASP-93 AND THR-112.
 RX MEDLINE=95038761; PubMed=7951252;
 RA Esteban J., Rosen D.R., Bowling A.C., Sapp P., McKenna-Yasek D.,
 RA O'Regan J.P., Beal M.F., Horvitz H.R., Brown R.H. Jr.;
 RA "Identification of two novel mutations and a new polymorphism in the
 RT gene for Cu/Zn superoxide dismutase in patients with amyotrophic
 RT lateral sclerosis.";
 RL Hum. Mol. Genet. 3:997-998(1994).
 RN [19]
 RP VARIANT FALS GLY-115.
 RX MEDLINE=95187174; PubMed=7881433;
 RA Kostzewa M., Burck-Lehmann U., Mueller U.;
 RA "Autosomal dominant amyotrophic lateral sclerosis: a novel mutation
 RT in the Cu/Zn superoxide dismutase-1 gene.";
 RL Hum. Mol. Genet. 3:2261-2262(1994).
 RN [20]
 RP VARIANT FALS ARG-46.
 RX MEDLINE=95138736; PubMed=7836951;
 RA Aoki M., Ogasawara M., Matsubara Y., Narisawa K., Nakamura S.,
 RA Itoyama Y., Abe K.;
 RA "Familial amyotrophic lateral sclerosis (ALS) in Japan associated
 RT with H46R mutation in Cu/Zn superoxide dismutase gene: a possible new
 RT subtype of familial ALS.";
 RL J. Neurol. Sci. 126:77-83(1994).
 RN [21]
 RP VARIANT FALS THR-113.
 RX MEDLINE=95089537; PubMed=7997024;
 RA Suthers G., Laing N., Wilton S., Dorosz S., Waddy H.;
 RA "Sporadic motoneuron disease due to familial SOD1 mutation with low
 RT penetrance.";
 RL Lancet 344:1773-1773(1994).
 RN [22]
 RP VARIANT FALS ASN-101.
 RX MEDLINE=95174778; PubMed=7870076;
 RA Jones C.T., Shaw P.J., Chari G., Brock D.J.;
 RA "Identification of a novel exon 4 SOD1 mutation in a sporadic
 RT amyotrophic lateral sclerosis patient.";
 RL Mol. Cell. Probes 8:329-330(1994).
 RN [23]
 RP VARIANTS FALS.
 RX MEDLINE=95193785; PubMed=7887412;
 RA Pramatarova A., Figlewicz D.A., Krizus A., Han F.Y.,
 RA Ceballos-Picot I., Nicole A., Dib M., Melinger V., Brown R.H.,
 RA Rouleau G.A.;
 RA "Identification of new mutations in the Cu/Zn superoxide dismutase
 RT gene of patients with familial amyotrophic lateral sclerosis.";
 RL Am. J. Hum. Genet. 56:592-596(1995).
 RN [24]
 RP VARIANT FALS ILB-148.

Query Match 99.4%; Score 816; DB 1; Length 153;
 Best Local Similarity 100.0%; Pred. No. 1e-68;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVFEGDNTAGCTSA 61
 |||||
 DB 1 ATKAVCVLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVFEGDNTAGCTSA 60
 |||||

QY 62 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVH 121
 |||||
 DB 61 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVH 120
 |||||

QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 |||||
 DB 121 EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 153
 |||||

RESULT 2
 SODC_CAVPO STANDARD; PRT; 152 AA.
 AC P33431;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 ON NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hartley; TISSUE=Lung;
 RA Yuan H.T., Bingle C.D., Kelly F.J.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=94059414; PubMed=8240718;
 RA Wolf B., Reinecke K., Aumann K.-D., Brigelius-Flohe R., Flohe L.;
 RT "Taxonomical classification of the guinea pig based on its Cu/Zn
 superoxide dismutase sequence."
 RL Biol. Chem. Hoppe-Seyler 374:641-649(1993).
 CC -1- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -1- COFACTOR: Copper and zinc.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U39844; AAC52720.1; -
 CC DR HSP: P00441; 4SOD.
 CC DR InterPro: IPR001424; SOD_CU_ZN.
 CC DR Pfam: PF00080; sodcu; 1.
 CC DR PRINTS: P00068; CUZNDISMUTASE.
 CC DR PRODOM: PD000469; SOD_CU_ZN; 1.
 CC DR PROSITE: PS00087; SOD_CU_ZN; 1.
 CC DR PROSITE: PS00332; SOD_CU_ZN; 2; 1.
 CC KW Oxidoreductase; Copper; Zinc; Acetylation.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT METAL 45 45 COPPER (BY SIMILARITY).

FT METAL 47 47 COPPER (BY SIMILARITY).
 FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 70 70 ZINC (BY SIMILARITY).
 FT METAL 79 79 ZINC (BY SIMILARITY).
 FT METAL 82 82 ZINC (BY SIMILARITY).
 FT METAL 119 119 COPPER (BY SIMILARITY).
 FT DISULFID 56 145 BY SIMILARITY.
 FT CONFLICT 102 103 LI -> IL (IN REF. 2).
 SQ SEQUENCE 152 AA; 15538 MW; 5891CE16E2ACBBF3 CRC64;

Query Match 83.4%; Score 684.5; DB 1; Length 152;
 Best Local Similarity 86.3%; Pred. No. 1.5e-56;
 Matches 132; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVFEGDNTAGCTSA 61
 |||||
 DB 1 ATKAVCVLKGDGPVQGIINFEQK-ANGPVVYKGRITGLVEGKHGFHVFEGDNTAGCTSA 59
 |||||

QY 62 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVH 121
 |||||
 DB 60 GPHEFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGSANSTIGRTMVVH 119
 |||||

QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 |||||
 DB 120 EXPDDLKGGNEESTKTGNAGSLACGVIGIAQ 152
 |||||

RESULT 3
 SODC_MOUSE STANDARD; PRT; 153 AA.
 AC P08228;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWR/J; TISSUE=Liver;
 RX MEDLINE=88203220; PubMed=3362683;
 RA Bewley G.C.;
 RT "cDNA and deduced amino acid sequence of murine Cu-Zn superoxide
 RT dismutase."
 RL Nucleic Acids Res. 16:2728-2728(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91216458; PubMed=2022332;
 RA Benedetto M.T., Anzai Y., Gordon J.W.;
 RT "Isolation and analysis of the mouse genomic sequence encoding
 RT Cu(2+)-Zn2+ superoxide dismutase."
 RL Gene 99:191-195(1991).
 RN [3]
 RP SEQUENCE OF 3-22.
 RX MEDLINE=90361747; PubMed=2391363;
 RA Pluthero F.G., Shreeve M., Eskinazi D., van der Gaag H., Huang K.S.,
 RA Hulmes J.D., Blum M., Axelrad A.A.;
 RT "Purification of an inhibitor of erythroid progenitor cell cycling
 RT and antagonist to interleukin 3 from mouse marrow cell supernatants
 RT and its identification as cytosolic superoxide dismutase."
 RL J. Cell Biol. 111:1217-1223(1990).
 CC -1- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -1- COFACTOR: Copper and zinc.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```

DR PIR; JC2015; JC2015.
DR HSSP; P00442; ICBJ.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc.
FT INIT_MET 0
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 48 48 COPPER (BY SIMILARITY).
FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 80 80 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
FT METAL 120 120 COPPER (BY SIMILARITY).
FT DISULFID 57 146 BY SIMILARITY.
SQ SEQUENCE 153 AA; 15780 MW; 1EF987CE7A0F79A0 CRC64;

Query Match 82.0%; Score 673; DB 1; Length 153;
Best Local Similarity 83.0%; Pred. No. 1.7e-55;
Matches 127; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
Db 1 AMKAVCVLKGDPVQGIINFEQKASGEVVSQITGLTEGEHGFHVGQDNTQGCCTA 60
QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHIIIGRTLWVH 121
Db 61 GPHFNPHSKKHGPKADEERHVGDLGNVAAGKGVADVSTEDSVISLSDHCHIIIGRTLWVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 EKQDDLKGGNEESTKTGNAGSRLACGVIGIAQ 153

RESULT 5
SODC_HORSE
ID SODC_HORSE STANDARD; PRT; 153 AA.
AC P00443;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN NCBI_TaxID=9796;
RP SEQUENCE FROM N.A.
RX MEDLINE=97080551; PubMed=8921896;
RA de la Rue-Domenech R., Wiedmann M., Mohammed H.O., Cummings J.F.,
RA Divers T.J., Batt C.A.;
RT "Equine motor neuron disease is not linked to Cu/Zn superoxide
RT dismutase mutations: sequence analysis of the equine Cu/Zn superoxide
RT dismutase cDNA."
RL Gene 178:83-88(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RC MEDLINE=99261591; PubMed=1031206;
RA Ishida N., Katayama Y., Sato F., Hasegawa T., Mukoyama H.;
RT "The cDNA sequences of equine antioxidant enzyme genes Cu/Zn-SOD and
RT Mn-SOD, and these expressions in equine tissues."
RL J. Vet. Med. Sci. 61:291-294(1999).
RN [3]
RP SEQUENCE.
RX MEDLINE=82052979; PubMed=7298616;
RA Lerch K., Ammer D.;
RT "Amino acid sequence of copper-zinc superoxide dismutase from horse
RT liver."

```

```

RL J. Biol. Chem. 256:11545-11551(1981).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -!- COFACTOR: Copper and zinc.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38956; AAC48682.1; -.
CC EMBL; AB001692; BAA76921.1; -.
CC PIR; A00515; DSHOCZ.
CC HSSP; P00442; ICBJ.
CC InterPro; IPR001424; SOD_CU_ZN.
CC Pfam; PF00080; sodcu; 1.
CC PRINTS; PR00068; CUZNDISMTASE.
CC ProDom; PD000469; SOD_CU_ZN; 1.
CC PROSITE; PS00087; SOD_CU_ZN_1; 1.
CC PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 48 48 COPPER (BY SIMILARITY).
FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
FT METAL 71 71 ZINC (BY SIMILARITY).
FT METAL 80 80 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
FT METAL 120 120 COPPER (BY SIMILARITY).
FT DISULFID 57 146 BY SIMILARITY.
SQ SEQUENCE 153 AA; 15940 MW; 202A190D0EFA0A43 CRC64;

Query Match 80.6%; Score 662; DB 1; Length 153;
Best Local Similarity 80.9%; Pred. No. 1.8e-54;
Matches 123; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
Db 1 ALKAVCVLKGDPVHGVTHFEQQEQGGPVVLKGFIEGLTKGDHGFHVEFGDNTQGCCTA 60
QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDHCHIIIGRTLWVH 121
Db 61 GAHFNPLSRKHGPKDEERHVGDLGNVTADENKADVDKDSVLSLKGKHSIIGRTMVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIA 153
Db 121 EKQDDLKGGNEESTKTGNAGSRLACGVIGIA 152

RESULT 6
SODC_RABBIT
ID SODC_RABBIT STANDARD; PRT; 152 AA.
AC P09212;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Lung;

```



```

RA Jackson R.M., Ho Y.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE.
RX MEDLINE=89105317; PubMed=3214553;
RA Reinecke K., Wolf B., Michelson A.M., Puget K., Steffens G.J.,
RA Flohe L.;
RL "The amino-acid sequence of rabbit Cu-Zn superoxide dismutase.";
RL Biol. Chem. Hoppe-Seyler 369:715-725(1988).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z22644; CAAB0357.1; -
DR PIR; S01134; S01134.
DR PIR; S33162; S33162.
DR HSP; P00441; 4SOD.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR PRODOM; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; FALSE_NEG.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; zinc; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT METAL 45 45 COPPER (BY SIMILARITY).
FT METAL 47 47 COPPER (BY SIMILARITY).
FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
FT METAL 70 70 ZINC (BY SIMILARITY).
FT METAL 79 79 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 119 119 COPPER (BY SIMILARITY).
FT METAL 145 145 BY SIMILARITY.
FT DISULFID 56 16 A -> G (IN REF. 2).
FT CONFLICT 16 16 E -> G (IN REF. 2).
FT CONFLICT 43 43 R -> T (IN REF. 2).
FT CONFLICT 53 53 S -> A (IN REF. 2).
FT CONFLICT 151 151
SQ SEQUENCE 152 AA; 15687 MW; 501415FFCB95571D CRC64;

Query Match 79.7%; Score 654.5; DB 1; Length 152;
Best Local Similarity 82.9%; Pred. No. 8.7e-54;
Matches 126; Conservative 8; Mismatches 17; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 59
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDVISLSGDHCHIGRTLVVH 121
DB 60 GPHFNPLSRKHGGPKDEERHVGDLGNVTAGSNVADVLIEDSVISLSDGMSVIGRTLVVH 119
QY 122 EKADDLCKGNGNEESTKTGNAGSLRACGVIGIA 153
DB 120 EKEDDLCKGNGNEESTKTGNAGSLRACGVIGIS 151

RESULT 7
SODC_CEREL STANDARD; PRT; 151 AA.
ID SODC_CEREL
AC 046412; 046413;

```

```

DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-JUN-1992 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=North American wapiti, and European red deer;
RX MEDLINE=98005640; PubMed=9345726;
RA He K., Wilton S.D., Tate M.L., Murphy M.P.;
RT "Characterization of the erythrocyte superoxide dismutase allozymes
RT in the deer Cervus elaphus.";
RL Anim. Genet. 28:299-301(1997).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U93268; AAB88115.1; -
DR EMBL; U93269; AAB88116.1; -
DR HSP; P00442; ICBJ.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR PRODOM; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; FALSE_NEG.
KW Oxidoreductase; Copper; zinc; Acetylation; Polymorphism.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT METAL 44 44 COPPER (BY SIMILARITY).
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 61 61 COPPER (BY SIMILARITY).
FT METAL 69 69 COPPER AND ZINC (BY SIMILARITY).
FT METAL 78 78 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
FT METAL 118 118 COPPER (BY SIMILARITY).
FT DISULFID 55 144 BY SIMILARITY.
FT VARIANT 8 8 L -> M (IN EUROPEAN RED DEER).
FT VARIANT 25 25 H -> N (IN EUROPEAN RED DEER).
SQ SEQUENCE 151 AA; 15697 MW; E718D474D3A22B57 CRC64;

Query Match 79.3%; Score 651; DB 1; Length 151;
Best Local Similarity 82.4%; Pred. No. 1.8e-53;
Matches 126; Conservative 8; Mismatches 17; Indels 2; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 58
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDVISLSGDHCHIGRTLVVH 121
DB 59 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDVISLSGDHCHIGRTLVVH 118
QY 122 EKADDLCKGNGNEESTKTGNAGSLRACGVIGIA 154
DB 119 EKPDLLGRGNGNEESTKTGNAGSLRACGVIGIA 151

```



```

RESULT 8
SODC_PIG
ID SODC_PIG STANDARD; PRT; 152 AA.
AC P04178;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=85231234; PubMed=3891411;
RA Schinina M.E., Barra D., Simmaco M., Bossa F., Rotilio G.;
RT "Primary structure of porcine Cu,Zn superoxide dismutase.";
RL FEBS Lett. 186:267-270(1985).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
DR PIR; A0514; DSPGZ.
DR HSP; P00442; ICBU.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom: PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; Zinc; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT METAL 45 45 COPPER (BY SIMILARITY).
FT METAL 47 47 COPPER (BY SIMILARITY).
FT METAL 62 62 COPPER AND ZINC (BY SIMILARITY).
FT METAL 70 70 ZINC (BY SIMILARITY).
FT METAL 79 79 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 119 119 COPPER (BY SIMILARITY).
FT DISULFID 56 145
FT ACT_SITE 142 142 ESSENTIAL ARGININE.
FT SEQUENCE 152 AA; 15761 MW; E60BF8C3446DEACC CRC64;

Query Match 79.2%; Score 650.5; DB 1; Length 152;
Best Local Similarity 82.4%; Pred. No. 2e-53;
Matches 126; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

QY 2 ATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFVHFEGDNTAGCTSA 61
Db 1 ATKAVCVLKGDPVQGIITVFELK-GEKTVLVGTGKGLAEGDGHGFVHFQGDNTQCTSA 59
QY 62 GPHFNPFSRKHGKDPDEERHVGDLGNVTADKDGADVSTEDSVLSGDHCIIIGRTLVVH 121
Db 60 GPHFNPFSRKHGKDPQERHVGDLGNVTAGKDGATVYIEDSVIALSGDHSIIGRTVMVH 119
QY 122 EKADLLGKGGNEESTKTGNAGSRACGVIGIAQ 154
Db 120 EKPPDLLGRGNEESTKTGNAGSRACGVIGITQ 152

RESULT 9
SODC_BOVIN
ID SODC_BOVIN STANDARD; PRT; 151 AA.
AC P00442;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
GN SOD1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91088332; PubMed=2263495;
RA Gibbs L.S., Shaffer J.B.;
RT "Nucleotide sequence of bovine copper/zinc superoxide dismutase cDNA
RT generated by the polymerase chain reaction.";
RL Nucleic Acids Res. 18:7171-7171(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Hallewell R.A., Imlay K.C., Lee P., Fong N.M., Gallegos C.,
RA Getzoff E.D., Tainer J.A., Cabelli D.E., Teekamp-Olson P.;
RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE.
RX MEDLINE=75060425; PubMed=4279916;
RA Steinman H.M., Naik V.R., Abernethy J.L., Hill R.L.;
RT "Bovine erythrocyte superoxide dismutase. Complete amino acid
RT sequence.";
RL J. Biol. Chem. 249:7326-7338(1974).
RN [4]
RP DISULFIDE BOND.
RX MEDLINE=75060426; PubMed=4436313;
RA Abernethy J.L., Steinman H.M., Hill R.L.;
RT "Bovine erythrocyte superoxide dismutase. Subunit structure and
RT sequence location of the intrasubunit disulfide bond.";
RL J. Biol. Chem. 249:7339-7347(1974).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=83085795; PubMed=7175933;
RA Tainer J.A., Getzoff E.D., Beem K.M., Richardson J.S.,
RA Richardson D.C.;
RT "Determination and analysis of the 2 A-structure of copper, zinc
RT superoxide dismutase.";
RL J. Mol. Biol. 160:181-217(1982).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=75158289; PubMed=1055410;
RA Richardson J.S., Thomas K.A., Rubin B.H., Richardson D.C.;
RT "Crystal structure of bovine Cu,Zn superoxide dismutase at 3-A
RT resolution: chain tracing and metal ligands.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:1349-1353(1975).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=92318263; PubMed=1619651;
RA Djinovic K., Coda A., Antolini L., Pelosi G., Desideri A., Falconi M.,
RA Rotilio G., Bolognesi M.;
RT "Crystal structure solution and refinement of the semisynthetic
RT cobalt-substituted bovine erythrocyte superoxide dismutase at 2.0-A
RT resolution.";
RL J. Mol. Biol. 226:227-238(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95371126; PubMed=7643403;
RA Rypniewski W.R., Mangani S., Bruni B., Orioli P.L., Casati M.,
RA Wilson K.S.;
RT "Crystal structure of reduced bovine erythrocyte superoxide dismutase
RT at 1.9-A resolution.";
RL J. Mol. Biol. 251:282-296(1995).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=99194860; PubMed=10092461;
RA Hough M.A., Hasnain S.S.;
RT "Crystallographic structures of bovine copper-zinc superoxide
RT dismutase reveal asymmetry in two subunits: functionally important
RT three and five coordinate copper sites captured in the same crystal.";
RL J. Mol. Biol. 287:579-592(1999).

```


[10]
 RN STRUCTURE AND MECHANISM.
 RX MEDLINE=84068148; PubMed=6316150;
 RA Tainer J.A., Getzoff E.D., Richardson J.S., Richardson D.C.;
 RT "Structure and mechanism of copper, zinc superoxide dismutase."; Nature 306:284-287(1983).
 RL Nature 306:284-287(1983).
 CC -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -!- COFACTOR: Copper and zinc.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -!- DATABASE: NAMP-Worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/S/SOD.html".
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X54799; -; NOT_ANNOTATED_CDS.
 DR EMBL; M81129; AAA73164.1; -.
 DR PIR; A00513; DSB0CZ.
 DR PIR; S13674; S13674.
 DR PDB; 2SOD; 15-JAN-91.
 DR PDB; 3SOD; 15-APR-93.
 DR PDB; 1COB; 31-OCT-93.
 DR PDB; 1SDA; 30-APR-94.
 DR PDB; 1SXA; 03-JUN-95.
 DR PDB; 1SXB; 03-JUN-95.
 DR PDB; 1SXC; 03-JUN-95.
 DR PDB; 1SXN; 18-MAR-98.
 DR PDB; 1SXS; 30-SEP-98.
 DR PDB; 1SXX; 30-SEP-98.
 DR PDB; 1CB4; 29-MAR-99.
 DR PDB; 1CBT; 29-MAR-99.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISMUTASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc; Acetylation; 3D-structure.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT METAL 44 44 COPPER.
 FT METAL 46 46 COPPER.
 FT METAL 61 61 COPPER AND ZINC.
 FT METAL 69 69 ZINC.
 FT METAL 78 78 ZINC.
 FT METAL 81 81 ZINC.
 FT METAL 118 118 COPPER.
 FT DISULFID 55 144
 FT STRAND 3 9
 FT STRAND 15 23
 FT TURN 24 25
 FT STRAND 26 34
 FT STRAND 39 46
 FT TURN 52 55
 FT HELIX 56 58
 FT STRAND 61 61
 FT TURN 64 65
 FT TURN 72 73
 FT TURN 79 80
 FT STRAND 81 87
 FT TURN 89 90
 FT STRAND 92 99
 FT STRAND 103 103
 FT STRAND 109 109

FT TURN 111 112
 FT STRAND 114 118
 FT TURN 124 125
 FT HELIX 132 135
 FT TURN 136 137
 FT STRAND 141 146
 FT STRAND 148 149
 SQ SEQUENCE 151 AA; 15551 MW; F9BACA8AE48AC867 CRC64;
 Query Match 78.8%; Score 647; DB 1; Length 151;
 Best Local Similarity 81.7%; Pred. No. 4.3e-53;
 Matches 125; Conservative 10; Mismatches 16; Indels 2; Gaps 1;
 Qy 2 ATKAVCLVKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSA 61
 Db 1 ATKAVCLVKGDPVQGTIHFEAK--GDTVVVTGSIITGLTEGDFHVFQGDNTQGCTSA 58
 Qy 62 GPHNPLSRKKGKPRDHRHVGDLGNVTADKGVADVSDVSLSGDHCIGRTLVVH 121
 Db 59 GPHNPLSKKGKPRDHRHVGDLGNVTADKGNVAIVDPLSLSGEYSIIIGRTVVVH 118
 Qy 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 Db 119 EKPDGLRGGNEESTKTGNAGSLACGVIGIAK 151
 RESULT 10
 SODC_SHEEP
 ID SODC_SHEEP STANDARD; PRT; 151 AA.
 AC P09670;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RX [1]
 RP SEQUENCE.
 RX MEDLINE=87030907; PubMed=3770194;
 RA Schinina M.E., Barra D., Gentilomo S., Bossa F., Capo C.,
 RA Rotilio G., Calabrese L.;
 RT "Primary structure of a cationic Cu,Zn superoxide dismutase. The sheep enzyme."; FEBS Lett. 207:7-10(1986).
 RL -!- FUNCTION: Destroys radicals which are normally produced within the cells and which are toxic to biological systems.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -!- COFACTOR: Copper and zinc.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR PIR; A24475; A24475.
 DR HSSP; P00442; 1CBJ.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISMUTASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc.
 FT METAL 44 44 COPPER (BY SIMILARITY).
 FT METAL 46 46 COPPER (BY SIMILARITY).
 FT METAL 61 61 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 69 69 ZINC (BY SIMILARITY).
 FT METAL 78 78 ZINC (BY SIMILARITY).
 FT METAL 81 81 ZINC (BY SIMILARITY).
 FT METAL 118 118 COPPER (BY SIMILARITY).
 FT DISULFID 55 144
 FT ACT_SITE 141 141 ESSENTIAL ARGININE.

SQ SEQUENCE 151 AA; 15563 MW; EDF854EB2B17DD98 CRC64;
 Query Match 77.3%; Score 635; DB 1; Length 151;
 Best Local Similarity 80.9%; Pred. No. 5.5e-52;
 Matches 123; Conservative 8; Mismatches 19; Indels 2; Gaps 1;
 QY 2 ATKAVCLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 61
 DB 1 ATKAVCLKGDGPVQGTIRFEAK--GDKVVVTGSIKGLTEGLHGFHVFHFGDNTAGCTSA 58
 QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 121
 DB 59 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVH 118
 QY 122 ERADDLGKGNNESTKTGNAGSLACGVIGIA 153
 DB 119 ERPDLLGRGNEESTKTGNAGSLACGVIGIA 150
 RESULT 11
 SODC_CHICK
 ID SODC_CHICK STANDARD; PRT; 153 AA.
 AC P80566; Q92059;
 DT 01-REB-1996 (Rel. 33, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 GN SOD1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spinal cord;
 RX MEDLINE=97142141; PubMed=8988375;
 RA Stanton J.L., Wilton S.D., Laing N.G.;
 RT "Characterisation of the chicken Cu,Zn superoxide dismutase gene.";
 RL DNA Seq. 6:357-360(1996).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Erythrocyte;
 RX MEDLINE=96215439; PubMed=8647082;
 RA Schinina M.E., Carlini P., Polticelli F., Zappacosta F., Bossa F.,
 RA Calabrese L.;
 RT "Amino acid sequence of chicken Cu, Zn-containing superoxide
 dismutase and identification of glutathionyl adducts at exposed
 cysteine residues.";
 RL Eur. J. Biochem. 237:433-439(1996).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- MASS SPECTROMETRY: MW=15600; MW_ERR=2; METHOD=Electrospray.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U28407; AAB88059.1; .
 CC HSP: P00441; LSPD.
 CC InterPro: IPR001424; SOD_CU_ZN.
 CC Pfam: PF00080; sodcu; 1.
 CC PRINTS: PR00068; CUZNDISMUTASE.
 CC ProDom: PD000469; SOD_CU_ZN; 1.
 CC PROSITE: PS00332; SOD_CU_ZN_2; 1.
 CC OXIDOREDUCTASE: Copper; Zinc; Repeat.
 CC MOD_RES 1 1 BLOCKED.

DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc; Acetylation.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT METAL 46 46 COPPER (BY SIMILARITY).
 FT METAL 48 48 COPPER (BY SIMILARITY).
 FT METAL 63 63 COPPER AND ZINC (BY SIMILARITY).
 FT METAL 71 71 ZINC (BY SIMILARITY).
 FT METAL 80 80 ZINC (BY SIMILARITY).
 FT METAL 83 83 ZINC (BY SIMILARITY).
 FT METAL 119 119 COPPER (BY SIMILARITY).
 FT METAL 145 145 BY SIMILARITY.
 FT DISULFID 153 153 WITH GLUTATHIONE.
 FT CONFLICT 13 13
 FT SEQUENCE 153 AA; 15572 MW; 594F5242EF2AFBD7 CRC64;
 Query Match 72.2%; Score 593; DB 1; Length 153;
 Best Local Similarity 73.5%; Pred. No. 4.3e-48;
 Matches 111; Conservative 19; Mismatches 19; Indels 2; Gaps 2;
 QY 4 KAVCVLKGDGPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVFHFGDNTAGCTSA 63
 DB 4 KAVCVMKGDADPEVGIHFQQ--QGSGPVKVTGKITGLSDGDFHVFHFGDNTAGCTSA 62
 QY 64 HFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGDHCHIIIGRTLVLVHEK 123
 DB 63 HFNPEKQHGPKDADRHVGDLGNVTA--KGGVAEVEIEDSVISLTPGPHCIIGRTMWHAK 121
 QY 124 ADDLGKGNNEESTKTGNAGSLACGVIGIA 154
 DB 122 SDDLGRGDNESKLTGNAGPRLACGVIGIA 152
 RESULT 12
 SODC_CARCR
 ID SODC_CARCR STANDARD; PRT; 166 AA.
 AC P80174;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
 OS Caretta caretta (Loggerhead).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Testudines; Cryptodira; Chelonioidae; Cheloniidae; Caretta.
 OX NCBI_TaxID=8467;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=93170321; PubMed=8436140;
 RA Schinina M.E., Bossa F., Lania A., Capo C.R., Carlini P.,
 RA Calabrese L.;
 RT "The primary structure of turtle Cu,Zn superoxide dismutase.
 Structural and functional irrelevance of an insert conferring
 proteolytic susceptibility.";
 RL Eur. J. Biochem. 211:843-849(1993).
 CC -|- FUNCTION: Destroys radicals which are normally produced within the
 cells and which are toxic to biological systems.
 CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -|- COFACTOR: Copper and zinc.
 CC -|- SUBUNIT: HOMODIMER.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR PIR: S29782; S29782.
 DR HSP: P00441; 4SOD.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR Pfam: PF00080; sodcu; 1.
 DR PRINTS: PR00068; CUZNDISMUTASE.
 DR ProDom: PD000469; SOD_CU_ZN; 1.
 DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
 KW Oxidoreductase; Copper; Zinc; Repeat.
 FT MOD_RES 1 1 BLOCKED.

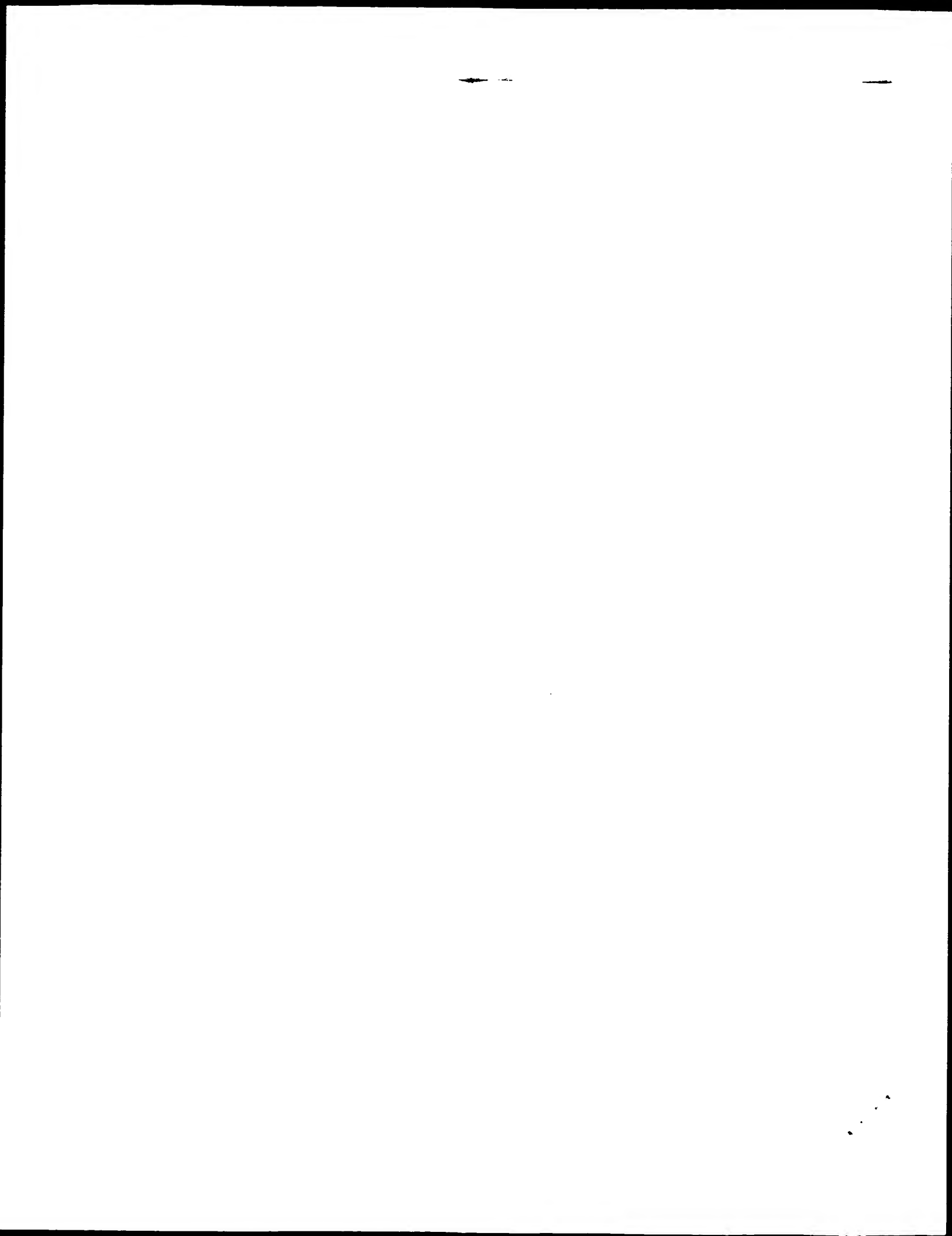
	ProDom; PD000469; SOD_CU_ZN_1.
DR	PROSITE; PS00087; SOD_CU_ZN_1; 1.
DR	PROSITE; PS00332; SOD_CU_ZN_2; 1.
KW	Oxidoreductase; Copper; Zinc.
FT	METAL 47 47 COPPER (BY SIMILARITY).
FT	METAL 49 49 COPPER (BY SIMILARITY).
FT	METAL 64 64 COPPER AND ZINC (BY SIMILARITY).
FT	METAL 72 72 ZINC (BY SIMILARITY).
FT	METAL 81 81 ZINC (BY SIMILARITY).
FT	METAL 84 84 ZINC (BY SIMILARITY).
FT	METAL 121 121 COPPER (BY SIMILARITY).
FT	DISULFID 58 147 BY SIMILARITY.
SQ	SEQUENCE 154 AA; 15953 MW; 62FA44CAI46C5057C CRG64;
Query Match 70.5%; Score 579; DB 1; Length 154;	
Best Local Similarity 70.1%; Pred.No. 8 6e-47;	
Matches 108; Conservative 10; Mismatches 36; Indels 0; Gaps	
QY	1 MATKAVCVLKGDPVGGIIINFEQKESGNPVKVMGSIKGLTEGLHGFHVHFEGDNTAGCTS 60
DQ	I I I I I I I I I I : : : : I
DB	1 MVNAKCVCVLKGTGEVTGVVFNGEKKPKVKTGITGLTTPGHKFPHVAFGDNTGCCIS 60
QY	61 AGHPNFSLRKHGCKPDERHRVGDLGNVTADKDGADVSDIESVISLGDHCIIGRTLVV 120
DQ	I I
DB	61 AGHPFNPDHTHGCTDSVRHHVDGLGNVTADASGAAKIEEDAMLTLSQHRSIGRTMTVI 120
QY	121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DQ	I I
DB	121 HEXEDDLGKGNEESLKTNAGGRLAGCGVIGITIQ 154
RESULT 14	
ID	SODC_PRIGL STANDARD; PRT; 152 AA.
AC	Pll418;
DT	01-OCT-1989 (Rel. 12, Created)
DT	01-OCT-1989 (Rel. 12, Last sequence update)
DE	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Superoxide dismutase [Cu-Zn] (EC 1.15.1.1).
OS	Priouace glauca (Blue shark).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC	Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC	Carcharinidae; Priouace.
OX	NCBI_TaxId=7815;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=89290032; PubMed=2500367;
RA	Calabrese L., Politicelli F., O'Neill P., Gaitieri A., Barra D.,
RA	Schinitz M.E., Bossa F.;
RT	"Substitution of arginine for lysine 134 alters electrostatic parameters of the active site in shark Cu,Zn superoxide dismutase."
RL	FEBB Lett. 250:49-52(1989);
CC	-!- FUNCTION: Destroys radicals which are normally produced within th e cells and which are toxic to biological systems.
CC	-!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC	-!- COFACTOR: Copper and zinc.
CC	-!- SUBUNIT: HOMODIMER.
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
PIR	S04623; S04623.
DR	PIR; S45643; S45643.
HSSP	P00442; ICBJ.
InterPro	IPR001424; SOD_CU_ZN.
PRam	PF00080; soccu; 1.
PRINTS	PR00068; CUZNDISMUTASE.
ProDom	PD000469; SOD_CU_ZN; 1.
PROSITE	PS00087; SOD_CU_ZN_1; 1.
PROSITE	PS00332; SOD_CU_ZN_2; 1.
Oxidoreductase; Copper; Zinc.	
Metal	44 44 COPPER.
FT METAL	46 46 COPPER.
FT METAL	61 61 COPPER AND ZINC.
FT METAL	69 69 ZINC.

FT METAL 78 78 ZINC.
FT METAL 81 81 ZINC.
FT METAL 118 118 COPPER.
FT DISULFID 55 144 BY SIMILARITY.
SQ SEQUENCE 152 AA; 15840 MW; 6617642A4F23C5AE CRC64;
Query Match 68.0%; Score 558.5; DB 1; Length 152;
Best Local Similarity 70.9%; Pred. No. 6.7e-45;
Matches 107; Conservative 15; Mismatches 28; Indels 1; Gaps 1;
QY 4 KAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVFEGDNTAGCTSAGP 63
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 2 KAVCVLKGTEVTGTPLEQ-AADGPTVLKSGITGLTPCKHGFHVFEGDNTAGCTSAGP 60
QY 64 HFNPLSRKHGPKDERHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 123
: : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 61 HYNPFKNGHGGPDDEHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 120
QY 124 ADDLKGKGGNEESTKTGNAGSRLACGVIGIAQ 154
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 121 EDDLKGGDEESLRTGNAGSRLACGVIGIAK 151

RESULT 15
SODD.XENLA STANDARD; PRT; 150 AA.
ID SODD.XENLA STANDARD; PRT; 150 AA.
AC P15107;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Cu-Zn] 2 (EC 1.15.1.1) (xSODB).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90221905; PubMed=2326205;
RA Carri M.T., Battistoni A., Mariottini P., Rotilio G.;
RT "Xenopus laevis Cu,Zn superoxide dismutase B cDNA sequence.";
RL Nucleic Acids Res. 18:1641-1641(1990).
RN [2]
RP SEQUENCE.
RX MEDLINE=89321563; PubMed=2751312;
RA Schinina M.E., Barra D., Bossa F., Calabrese L., Montesano L.,
RA Carri M.T., Mariottini P., Amaldi F., Rotilio G.;
RT "Primary structure from amino acid and cDNA sequences of two Cu,Zn
RT superoxide dismutase variants from Xenopus laevis.";
RL Arch. Biochem. Biophys. 272:507-515(1989).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=91376038; PubMed=1896428;
RA Falconi M., Rotilio G., Desideri A.;
RT "Modelling the three-dimensional structure and electrostatic
RT potential field of the two Cu,Zn superoxide dismutase variants from
RT Xenopus laevis.";
RL Proteins 10:149-155(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.49 ANGSTROMS).
RA Djinovic Carugo K., Battistoni A., Carri M.T., Polticelli F.,
RA Desideri A., Rotilio G., Coda A., Wilson K.S., Bolognesi M.;
RT "Three-dimensional structure of Xenopus laevis Cu,Zn superoxide
RT dismutase B. determined by X-ray crystallography at 1.5-A
RT resolution.";
RL Acta Crystallogr. D 52:176-188(1996).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Copper and zinc.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51518; CAA35890.1; -.
DR PIR: S05022; S05022.
DR PIR: S09568; S09568.
DR PDB: 1XSO; 10-JUL-95.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam: PF00080; sodcu; 1.
DR PRINTS: PRO0068; CUZNDISMUTASE.
DR PRODOM: PD000469; SOD_CU_ZN; 1.
DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
KW Oxidoreductase; Copper; zinc; 3D-structure.
FT INIT_MET 0 0
FT METAL 44 44 COPPER (BY SIMILARITY).
FT METAL 46 46 COPPER (BY SIMILARITY).
FT METAL 61 61 COPPER AND ZINC (BY SIMILARITY).
FT METAL 69 69 ZINC (BY SIMILARITY).
FT METAL 78 78 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
FT METAL 117 117 COPPER (BY SIMILARITY).
FT DISULFID 55 143 BY SIMILARITY.
FT CONFLICT 60 60 P -> S (IN REF. 1).
SQ SEQUENCE 150 AA; 15297 MW; 0F5236CEEB236250 CRC64;
Query Match 65.9%; Score 541; DB 1; Length 150;
Best Local Similarity 69.6%; Pred. No. 2.7e-43;
Matches 103; Conservative 16; Mismatches 27; Indels 2; Gaps 2;
QY 4 KAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVFEGDNTAGCTSAGP 63
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 2 KAVCVLAGSDGVKGVHFEQD-EGAVSVECKIEGLTDLGHGFHVFEGDNTAGCTSAGP 60
QY 64 HFNPLSRKHGPKDERHVGDLGNVTADKDGADVSDISGSDHCIIIGRTLVVHEK 123
|||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 61 HFNPNKNHGA PGDTDRHVGDLGNVTAE-GGVAQFKITDSLKLPNSIIIGRTAVVHEK 119
QY 124 ADDLKGKGGNEESTKTGNAGSRLACGVIG 151
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Db 120 ADDLKGKGGNEESTKTGNAGSRLACGVIG 147

Search completed: January 28, 2003, 14:15:18
Job time : 3.35564 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 7.78203 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rhodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	807	98.3	179	3	Q12728	Q12728 Schizosacch
2	638.5	77.8	153	6	Q8WNN6	Q8WNN6 canis famil
3	632.5	77.0	147	6	Q95ME5	Q95ME5 sus scrofa
4	563	68.6	154	13	Q8QHI0	Q8QHI0 oncorhynch
5	508.5	61.9	154	3	Q9C0N4	Q9C0N4 filobasidie
6	503.5	61.3	154	3	Q9C0T2	Q9C0T2 cryptococcu
7	500	60.9	153	5	Q9U4X3	Q9U4X3 drosophila
8	499	60.8	153	5	Q9U4X5	Q9U4X5 drosophila
9	496	60.4	153	5	Q9U4X2	Q9U4X2 drosophila
10	489.5	59.6	154	3	Q9C0S4	Q9C0S4 cryptococcu
11	489	59.6	145	5	Q96348	Q96348 drosophila
12	489	59.6	153	5	Q9U4X4	Q9U4X4 drosophila
13	479	58.3	158	10	Q9AR76	Q9AR76 populus tre
14	478.5	58.3	158	3	Q96WH8	Q96WH8 debaryomyce
15	477	58.1	145	5	O16861	O16861 chymomyza p
16	477	58.1	145	5	O96353	O96353 drosophila

17	476	58.0	145	5	O96349	O96349 drosophila
18	475	57.9	145	5	O16863	O16863 zaprionus t
19	473.5	57.7	183	3	O9C402	O9C402 cryptococcu
20	473.5	57.7	215	10	Q944B7	Q944B7 pinus pinas
21	472	57.5	158	10	Q9AR77	Q9AR77 populus tre
22	471.5	57.4	151	5	O77243	O77243 dictyosteli
23	471	57.4	145	5	O27630	O27630 drosophila
24	471	57.4	145	5	O8WRP5	O8WRP5 taenia soli
25	470	57.2	152	5	O96352	O96352 drosophila
26	468.5	57.1	210	8	O9BA07	O9BA07 populus tre
27	468	57.0	145	5	O96351	O96351 drosophila
28	467.5	56.9	152	10	Q9AR78	Q9AR78 populus tre
29	465.5	56.7	201	10	O24400	O24400 triticum ae
30	465	56.6	146	5	O27770	O27770 drosophila
31	464	56.5	145	5	O96356	O96356 drosophila
32	462.5	56.3	201	8	O96123	O96123 triticum ae
33	461	56.2	145	5	O16862	O16862 drosophila
34	461	56.2	146	5	O9XY94	O9XY94 fasciola he
35	460.5	56.1	201	5	O18483	O18483 onchocerca
36	460.5	56.1	201	10	O93Y61	O93Y61 dichanthelli
37	460	56.0	145	5	O27569	O27569 drosophila
38	460	56.0	158	5	O9Y0A5	O9Y0A5 acanthochei
39	459.5	56.0	156	10	Q9AR80	Q9AR80 pinus sylve
40	458.5	55.8	216	10	Q9SUJ7	Q9SUJ7 arabidopsis
41	458	55.8	153	3	O8X1S8	O8X1S8 aspergillus
42	454.5	55.4	152	10	O8S3V1	O8S3V1 sandersonia
43	453	55.2	161	10	O48654	O48654 marchantia
44	452.5	55.1	152	10	Q9ZS54	Q9ZS54 pinus sylve
45	452.5	55.1	152	10	P93801	P93801 zea mays (m

ALIGNMENTS

RESULT 1

Q12728 PRELIMINARY; PRT; 179 AA.
AC Q12728;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Yeast HB-SOD (Fragment).
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9138420; PubMed=1885572;
RA Inoue M., Watanabe N., Matsuno K., Sasaki J., Tanaka Y.,
RA Hatanaka-Sawai H., Amachi T.;
RT "Expression of a hybrid Cu/Zn-type superoxide dismutase which has high
RT affinity for heparin-like proteoglycans on vascular endothelial
RT cells.";
RL J. Biol. Chem. 266:16409-16414(1991).
DR EMBL; D90358; BAA14373.1; -;
DR HSSP; P00441; ISP.D.
DR InterPro; IPR001424; SOD_CU_2N.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PR00068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_2N; 1.
DR PROSITE; PS00087; SOD_CU_2N_1; 1.
DR PROSITE; PS00332; SOD_CU_2N_2; 1.
FT NON_TER 1 179
FT NON_TER 179 179
SQ SEQUENCE 179 AA; 18995 MW; 42D662A6785F6135 CRC64;

Query Match 98.3%; Score 807; DB 3; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.3e-69;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVGWSIKGLTEGLHGFHVHFGDNTAGCTSA 61


```

|||||
Db 1 ATRAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSA 60
Qy 62 GPHFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWVH 121
Db 61 GPHFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWVH 120
Qy 122 EKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 EKADDLGKGGNEESTKTGNAGSLACGVIGI 151

RESULT 2
Q8WNN6 PRELIMINARY; PRT; 153 AA.
AC Q8WNN6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cu/Zn superoxide dismutase.
GN SOD1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Green S.L., Tolwani R.J., Varma S., Quignon P., Galibert F.,
RA Cork L.C.;
RT "The canine Cu/Zn superoxide dismutase gene (SOD1).";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF346417; AAL61608.1; -.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PRINTS; PRO0068; CUZNDISMTASE.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00332; SOD_CU_ZN_2; UNKNOWN_1.
SQ SEQUENCE 153 AA; 15912 MW; 0D7900E59C57B6B0 CRC64;

Query Match 77.8%; Score 638.5; DB 6; Length 153;
Best Local Similarity 80.5%; Pred. No. 2.3e-53;
Matches 124; Conservative 11; Mismatches 18; Indels 1; Gaps 1;

Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MEMKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 59
Qy 61 AGPHFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWV 120
Db 60 AGPHFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWV 119
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 120 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 153

RESULT 3
Q95WE5 PRELIMINARY; PRT; 147 AA.
AC Q95WE5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Superoxide dismutase 1 (Fragment).
GN SOD1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Llu B.S., Ding S.T.;
RT "Expression of superoxide dismutase 1 in pigs.";

```

```

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF396674; AAR84037.1; -.
DR InterPro: IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR ProDom; PD000469; SOD_CU_ZN; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; UNKNOWN_1.
DR PROSITE; PS00332; SOD_CU_ZN_2; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 147
SQ SEQUENCE 147 AA; 15246 MW; 62156DD7EC5430A4 CRC64;

Query Match 77.0%; Score 632.5; DB 6; Length 147;
Best Local Similarity 82.4%; Pred. No. 8.3e-53;
Matches 122; Conservative 7; Mismatches 18; Indels 1; Gaps 1;

Qy 4 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSAGP 63
Db 1 KAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTSAGP 59
Qy 64 HFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWVHEK 123
Db 60 HFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWVHEK 119
Qy 124 ADDLGKGGNEESTKTGNAGSLACGVIGI 151
Db 120 PDDLGRGGNEESTKTGNAGSLACGVIGI 147

RESULT 4
Q8QHI0 PRELIMINARY; PRT; 154 AA.
AC Q8QHI0;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cu/Zn superoxide dismutase.
GN SOD1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Brunelli J.P., Thorgaard G.H.;
RT "Rainbow trout Cu/Zn-superoxide dismutase (SOD1) mRNA.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF469663; AAL79162.1; -.
SQ SEQUENCE 154 AA; 15755 MW; CA3C983D78067A18 CRC64;

Query Match 68.6%; Score 563; DB 13; Length 154;
Best Local Similarity 68.2%; Pred. No. 3.8e-46;
Matches 105; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

Qy 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Db 1 MAMKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
Qy 61 AGPHFNPLSRKHGKPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLWV 120
Db 61 AGPHFNPHNQTHGGPTDAVRHVGDLGNVTAGADNVAKINQDKMLTLGPDSSIIGRTWVI 120
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154

RESULT 5
Q9CON4 PRELIMINARY; PRT; 154 AA.
AC Q9CON4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```


DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Copper zinc superoxide dismutase.
 GN SOD1.
 OS Filobasidiella bacillisporea.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=37769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WM 780-C;
 RA Chaturvedi S., Chaturvedi V.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NYS 155B;
 RA Chaturvedi S., Hamilton A.J., Chaturvedi V.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NYS 155B;
 RA Hamilton A.J., Chaturvedi S., Chaturvedi V.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR EMBL: AF248051; AAK31920.1; -;
 DR EMBL: AF248049; AAK31918.1; -;
 DR EMBL: AF248045; AAK31919.1; -;
 DR HSP: P00442; 3SOD.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR Pfam: PF00080; sodcu; 1.
 DR PRINTS: PR00068; CUZNDISMTASE.
 DR PRODOM: PD000469; SOD_CU_ZN; 1.
 DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
 DR Copper; Oxidoreductase; Zinc.
 KW SEQUENCE 154 AA; 15637 MW; CM4C2B4BA8F43FE4 CRC64;
 SQ

Query Match 61.9%; Score 508.5; DB 3; Length 154;
 Best Local Similarity 64.9%; Pred. No. 6.2e-41;
 Matches 98; Conservative 11; Mismatches 41; Indels 1; Gaps 1;
 QY 4 KAVCVLKGDGPGVQGIINFEQKESNGPVKVGSIKGL-TEGLHGFHVFHFGDNTAGCTSGAG 62
 DB 3 KAVAVLKGDSPTGVTITFQREGAPVTGSDIKNLDAERGFHVFHFGDNTAGCTSGAG 62
 QY 63 PHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIKSLSGDHCITGRTLVVHA 122
 DB 63 PHFNPHKNGHGAPESSERHVGDLGNVTGNGGVAMVDISKVLSLFGPHSIIIGRSMVHA 122
 QY 123 KADDLGKGNEESTKTNAGSLACGVIGIA 153
 DB 123 GTDDFGKGGNAESLKTGNAGAACGVIGIS 153
 RESULT 6
 Q9COT2 PRELIMINARY; PRT; 154 AA.
 AC Q9COT2;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 20, Last annotation update)
 DE Copper zinc superoxide dismutase.
 GN SOD1.
 OS Cryptococcus neoformans var. neoformans.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=40410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B-3501;
 RA Chaturvedi S., Hamilton A.J., Chaturvedi V.;

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-B-3501;
 RA Chaturvedi S., Chaturvedi V.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR EMBL: AF248048; AAK31917.1; -;
 DR EMBL: AF248047; AAK31916.1; -;
 DR HSP: P00441; 4SOD.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR Pfam: PF00080; sodcu; 1.
 DR PRINTS: PR00068; CUZNDISMTASE.
 DR PRODOM: PD000469; SOD_CU_ZN; 1.
 DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
 DR Copper; Oxidoreductase; Zinc.
 KW SEQUENCE 154 AA; 15779 MW; OB0054E88FECB35 CRC64;
 SQ

Query Match 61.3%; Score 503.5; DB 3; Length 154;
 Best Local Similarity 66.9%; Pred. No. 1.9e-40;
 Matches 101; Conservative 7; Mismatches 42; Indels 1; Gaps 1;
 QY 4 KAVCVLKGDGPGVQGIINFEQKESNGPVKVGSIKGL-TEGLHGFHVFHFGDNTAGCTSGAG 62
 DB 3 KAVAVLKGDSHVYGTITFTQSEGAFCVSGEIKNLDAERGFHVFHFGDNTAGCTSGAG 62
 QY 63 PHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIKSLSGDHCITGRTLVVHE 122
 DB 63 PHFNPHKNGHGAPESSERHVGDLGNVTGNGGVAMVDISKVLSLFGPHSIIIGRSMVHA 122
 QY 123 KADDLGKGNEESTKTNAGSLACGVIGIA 153
 DB 123 GTDDLGKGGNEESLKTGNAGARLACGVIGIA 153
 RESULT 7
 Q9U4X3 PRELIMINARY; PRT; 153 AA.
 AC Q9U4X3;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE CU-ZN superoxide dismutase.
 GN SOD.
 OS Drosophila yakuba (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Arxontaki K., Kastanis P., Tsakas S., Loukas M., Ellipoulos E.;
 RT "Phylogenetic analysis of Drosophila melanogaster group based on Cu-Zn
 RT superoxide dismutase gene sequences."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
 CC -1- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR EMBL: AF127159; AAF23598.1; -;
 DR HSP: P00441; 4SOD.
 DR FlyBase: FBgn0022829; Dyak\Sod.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR Pfam: PF00080; sodcu; 1.
 DR PRINTS: PR00068; CUZNDISMTASE.
 DR PRODOM: PD000469; SOD_CU_ZN; 1.
 DR PROSITE: PS00087; SOD_CU_ZN_1; 1.
 DR PROSITE: PS00332; SOD_CU_ZN_2; 1.
 DR Copper; Oxidoreductase; Zinc.
 KW

SQ SEQUENCE 153 AA; 15744 MW; 569DB27D4C545D10 CRC64;
 Query Match 60.9%; Score 500; DB 5; Length 153;
 Best Local Similarity 62.3%; Pred. No. 4e-40;
 Matches 96; Conservative 17; Mismatches 39; Indels 2; Gaps 1;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 DB 1 MVYKAVCVINGD--AKGTVFVEQSESTPVKVSQEVGCLAKGLHGFHVFHFGDNTAGCTS 58
 QY 61 AGPHFNPLSRKHGPKDERHVGDLGNVTADKGVADVSIEDSVLSGDHCIIIGRTLVV 120
 DB 59 SGPHFNPKGKHGAPVDENRHLGDLGNIEATGDCPTKVSITDSRTILFGADSIIGRTVVV 118
 QY 121 HEKADDLGKGGNEESTKTGNAGSRACGVIGIAQ 154
 DB 119 HADADDLGGGHELKSKSTGNAGARICGCVIGIAK 152
 RESULT 8
 Q9U4X5 PRELIMINARY; PRT; 153 AA.
 AC Q9U4X5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE CU-ZN superoxide dismutase.
 OS Drosophila orena (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7233;
 RN [1]
 RA Arxontaki K., Tsakani P., Tsakas S., Loukas M., Eliopoulos E.;
 RP "Phylogenetic analysis of Drosophila melanogaster group based on Cu-Zn
 RT superoxide dismutase gene sequences."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR EMBL; AF127155; AAF23594.1; -.
 DR HSP; P00441; 4SOD.
 DR FlyBase; FBgn029310; Dore\Sod.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISMTASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN.1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN.2; 1.
 KW Copper; Oxidoreductase; Zinc.
 SQ SEQUENCE 153 AA; 15772 MW; 569DB26D1E295E0 CRC64;
 Query Match 60.8%; Score 499; DB 5; Length 153;
 Best Local Similarity 62.3%; Pred. No. 4.9e-40;
 Matches 96; Conservative 17; Mismatches 39; Indels 2; Gaps 1;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 DB 1 MVYKAVCVINGD--AKGTVFVEQSESTPVKVSQEVGCLAKGLHGFHVFHFGDNTAGCTS 58
 QY 61 AGPHFNPLSRKHGPKDERHVGDLGNVTADKGVADVSIEDSVLSGDHCIIIGRTLVV 120
 DB 59 SGPHFNPKGKHGAPVDENRHLGDLGNIEATGDCPTKVSITDSRTILFGADSIIGRTVVV 118
 QY 121 HEKADDLGKGGNEESTKTGNAGSRACGVIGIAQ 154
 DB 119 HADADDLGGGHELKSKSTGNAGARICGCVIGIAK 152
 RESULT 9

Q9U4X2 PRELIMINARY; PRT; 153 AA.
 AC Q9U4X2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE CU-ZN superoxide dismutase.
 OS Drosophila teissieri (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7243;
 RN [1]
 RA Arxontaki K., Tsakani P., Tsakas S., Loukas M., Eliopoulos E.;
 RP "Phylogenetic analysis of Drosophila melanogaster group based on Cu-Zn
 RT superoxide dismutase gene sequences."
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR EMBL; AF127160; AAF23599.1; -.
 DR HSP; P00441; 4SOD.
 DR FlyBase; FBgn0026098; Dtei\Sod.
 DR InterPro; IPR001424; SOD_CU_ZN.
 DR Pfam; PF00080; sodcu; 1.
 DR PRINTS; PR00068; CUZNDISMTASE.
 DR ProDom; PD000469; SOD_CU_ZN; 1.
 DR PROSITE; PS00087; SOD_CU_ZN.1; 1.
 DR PROSITE; PS00332; SOD_CU_ZN.2; 1.
 KW Copper; Oxidoreductase; Zinc.
 SQ SEQUENCE 153 AA; 15744 MW; 569DA9DCF6545D10 CRC64;
 Query Match 60.4%; Score 496; DB 5; Length 153;
 Best Local Similarity 61.7%; Pred. No. 9.6e-40;
 Matches 95; Conservative 18; Mismatches 39; Indels 2; Gaps 1;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 DB 1 MVYKAVCVINGD--AKGTVFVEQSESTPVKVSQEVGCLAKGLHGFHVFHFGDNTAGCTS 58
 QY 61 AGPHFNPLSRKHGPKDERHVGDLGNVTADKGVADVSIEDSVLSGDHCIIIGRTLVV 120
 DB 59 SGPHFNPKGKHGAPVDENRHLGDLGNIEATGDCPTKVSITDSRTILFGADSIIGRTVVV 118
 QY 121 HEKADDLGKGGNEESTKTGNAGSRACGVIGIAQ 154
 DB 119 HADADDLGGGHELKSKSTGNAGARICGCVIGIAK 152
 RESULT 10
 Q9C0S4 PRELIMINARY; PRT; 154 AA.
 AC Q9C0S4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Copper zinc superoxide dismutase.
 GN SOD1.
 OS Cryptococcus neoformans var. neoformans.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=40410;
 RN [1]
 RA Chaturvedi S., Hamilton A.J., Lowery C.V., Chaturvedi V.;
 RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H99;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H99;

QY 72 HGGPKDEERHVGDLGNWTADKDGADVSDSVISLSGDHCHIGRTLVVHERADDLGKGG 131
Db 62 HGAPADENRHLGLDGNIEAPCGDPTKVCINDCKITLFEHSIVGRTVVVHADPDDLKGG 121
QY 132 NEESTKTNAGSRLACGVIGIAQ 154
Db 122 HELSKSTGNAGARIGCGVIGICK 144

Search completed: January 28, 2003, 14:18:27
Job time : 8.78203 secs

XX
TJ

PS Disclosure; Figure 3; 25pp; German.
 XX
 CC cDNA encoding hSOD (AAN71204) is isolated using a 5-triplet probe
 CC (AAN71205). The cDNA is subjected to site-specific mutagenesis and
 CC expressed to produce hSOD analogues (AAP70930) which are claimed.
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 8; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 RESULT 2
 AAR61336
 ID AAR61336 standard; Protein; 154 AA.
 XX
 AC AAR61336;
 XX
 DT 31-MAY-1995 (first entry)
 XX
 DE Human superoxide dismutase SOD-1.
 XX
 KW Human superoxide dismutase; hSOD1; neurodegeneration;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW Hallervorden-Spatz disease; olivopontocerebellar atrophy;
 KW familial amyotrophic lateral sclerosis; FALS; diagnosis; mutant SOD;
 KW reverse transcription polymerase chain reaction; SSCP analysis.
 XX
 OS Homo sapiens.
 XX
 PN W09419493-A.
 XX
 PD 01-SEP-1994.
 XX
 PF 28-FEB-1994; 94WO-US02089.
 XX
 PR 26-FEB-1993; 93US-0023980.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Brown R, Horvitz HR, Rosen DR;
 XX
 WPI; 1994-294353/36.
 DR N-PSDB; AAQ67473.
 XX
 XX
 XX
 PT e.g. amyotrophic lateral sclerosis, which are the result of e.g.
 PT decreased SOD activity
 XX
 PS Claim 22; ; 94pp; English.
 XX
 CC The presence of a mutation in a gene encoding a superoxide
 CC dismutase (SOD1, SOD2 or SOD3) indicates an increased likelihood of
 CC developing a cell death disease, specifically a neurodegenerative
 CC disease. The use of SOD polypeptides to treat amyotrophic lateral
 CC sclerosis and diseases involving a deleterious mutation in the
 CC glutathione peroxidase-encoding gene, the catalase-encoding gene
 CC and the nitric oxide-encoding gene is claimed. The specification
 CC includes the sequences of human SOD1, hSOD2 and hSOD3 (AAR61336-
 CC

CC AAR61338, respectively).
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 15; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVLSGDHCHIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 RESULT 3
 AAR91931
 ID AAR91931 standard; Protein; 154 AA.
 XX
 AC AAR91931;
 XX
 DT 16-DEC-1996 (first entry)
 XX
 DE Human superoxide dismutase.
 XX
 KW Human; superoxide dismutase; hSOD; Saccharomyces cerevisiae;
 KW molasses; medium; recombinant production; microbiology;
 KW medicine.
 XX
 OS Homo sapiens.
 XX
 PN RU2044771-C1.
 XX
 PD 27-SEP-1995.
 XX
 PF 01-JUN-1993; 93RU-0028609.
 XX
 PR 01-JUN-1993; 93RU-0028609.
 XX
 PA (TRIS-) TRIS STOCK CO LTD.
 XX
 PI Efremov VD, Kozlov DG, Vinetskii YP;
 XX
 DR WPI; 1996-228796/23.
 DR N-PSDB; AAT29715.
 XX
 PT Human superoxide dismutase-producing Saccharomyces cerevisiae
 PT strain - can grow on medium which contains molasses and produces
 PT hSOD yields corresp. to 20 % of cell protein
 XX
 PS Disclosure; Columns 5-6; 6pp; Russian.
 XX
 CC The present sequence is human superoxide dismutase (hSOD), the cDNA
 CC of which was used in the prepn. of a Saccharomyces cerevisiae
 CC strain which can grow on a medium contg. molasses, and produces
 CC hSOD yields corresp. to 20 % of cell protein. The recombinant hSOD
 CC produced can be used in the microbiological, and medical
 CC industries.
 XX
 SQ Sequence 154 AA;
 Query Match 100.0%; Score 821; DB 17; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.1e-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60

Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4
 AAW96318
 ID AAW96318 standard; Protein; 154 AA.
 XX
 AC AAW96318;
 XX
 DT 28-JUN-1999 (first entry)
 DE
 XX Cu/Zn superoxide dismutase.
 XX Manganese containing superoxide dismutase; MnSOD; IDDM;
 KW diabetes mellitus; treatment; therapy; nitric oxide; NO; beta cell;
 KW fatty acid; lipotoxic; cytotoxic; cytokine; osteoporosis;
 KW inflammatory disease; autoimmune disease; neurodegenerative disease.
 XX Homo sapiens.
 OS
 XX WQ9906059-A2.
 PN
 XX 11-FEB-1999.
 PD
 XX 30-JUL-1998; 98MO-US15781.
 XX
 PF 03-MAR-1998; 98US-0055092.
 PR 30-JUL-1997; 97US-0055092.
 XX
 XX (BETA-) BETAGENE INC.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Clark SA, Hohmeier H, Koyama K, Lee Y, Newgard CB;
 PI Ohneda M, Shimabukuro, Thigpen A, Unger RH;
 XX
 DR WPI: 1999-153448/13.
 DR N-PSDB; AAX08430.
 XX
 PT Protection of mammalian cells against immunotoxicity or lipotoxicity
 PT - used for treating, e.g. diabetes, obesity, wasting syndromes,
 PT osteoporosis, inflammatory diseases, autoimmune diseases or
 PT neurodegenerative diseases
 XX
 PS Disclosure; Page 235-236; 253pp; English.
 XX

Inhibition of cytokine mediated immunotoxicity of cells can be
 achieved by blocking free radical production or the accumulation of
 free radicals in that cell. Treatment of insulin dependent diabetes
 mellitus (IDDM) can be achieved by blocking nitric oxide (NO)
 production in a pancreatic beta cell and by providing a composition
 comprising an agent that reduces levels of fatty acids in the cells
 and protects beta-cells of the subject against lipid-mediated cell
 death. Cells can also be protected against nitric oxide mediated
 cytotoxicity by introducing into the cell an antioxidant agent.
 The methods can be used for protecting cells against immunotoxicity
 mediated by, e.g. IL-1 beta, IL-1 alpha, gamma IFN, TNF alpha, TNF
 beta, IL-8, IL-6, IL-2, IL-3, IL-5, IL-7, IL-9, IL-14, IL-17,
 CC granulocyte-macrophage colony stimulating factor or monocyte
 CC chemoattractant protein-1. The methods can be used for the treatment
 CC of e.g. insulin-dependent diabetes mellitus (IDDM), NIDDM, obesity,
 CC wasting syndromes, short stature, osteoporosis, inflammatory
 CC diseases, autoimmune diseases, or neurodegenerative diseases.
 XX
 SQ Sequence 154 AA;

Query Match 100.0%; Score 821; DB 20; Length 154;
 Best Local Similarity 100.0%; Pred. No. 2.le-86;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVFHFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
 Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 5
 AAW10377
 ID AAW10377 standard; Protein; 154 AA.
 XX
 AC AAW10377;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Superoxide dismutase 1 (SOD1).
 XX
 KW Superoxide dismutase 1; soluble amyotrophic lateral sclerosis 1 (adult);
 KW haplotyping; SOD1.
 XX Homo sapiens.
 OS
 XX WO200185741-A2.
 PN
 XX 15-NOV-2001.
 PD
 XX 07-MAY-2001; 2001WO-US14772.
 PF
 XX 05-MAY-2000; 2000US-202491P.
 PR
 XX (GENA-) GENAISSANCE PHARM INC.
 PA
 PI Choi JY, Bentivegna SC, Kliem SE, Koshy B, Parks KE;
 PI N-PSDB; AAS16687.
 DR WPI: 2002-055578/07.
 DR N-PSDB; AAS16687.
 XX
 PT Isolated human superoxide dismutase 1 (SOD1) soluble polynucleotide,
 PT useful for screening therapeutic compounds, comprises a sequence which
 PT is a polymorphic variant of reference sequence for the SOD1 gene or its
 PT fragment -
 XX
 PS Disclosure; Fig 3; 70pp; English.
 XX

The invention relates to an isolated human superoxide dismutase 1,
 CC soluble (amyotrophic lateral sclerosis 1 (adult)) (SOD1) polynucleotide
 CC (1) comprising a sequence which is a polymorphic variant of a reference
 CC sequence for the SOD1 gene. Haplotyping the SOD1 gene of an individual,
 CC involves: (a) determining whether the individual has one of the SOD1
 CC haplotypes or haplotype pairs given in the specification; or
 CC (b) determining for one copy of the SOD1 gene present in the individual,
 CC the identity of the nucleotide at two or more polymorphic sites selected
 CC from PSI-7. The method is useful for determining whether an individual
 CC has a haplotype or haplotype pairs defined in the specification. The
 CC method is also useful for improving the efficacy and reliability of
 CC several steps in the discovery and development of drugs for treating
 CC diseases associated with SOD1 activity, e.g., amyotrophic lateral
 CC sclerosis, and to validate SOD1 as a candidate agent for treating a
 CC specific condition or disease associated with SOD1 activity. It can
 CC further be used in the design of clinical trials of candidate drugs for
 CC treating a specific condition or disease predicted to be associated with
 CC SOD1 activity. (1) is useful in studying the expression and function of
 CC SOD1, and in expressing SOD1 protein for use in screening for candidate

CC drugs to treat diseases related to SOD1 activity. The present sequence
CC represents the amino acid sequence of human superoxide dismutase 1,
CC soluble (amyotrophic lateral sclerosis 1 (adult)) (SOD1) as described in
CC the method of the invention.
XX

SQ Sequence 154 AA;
Query Match 100.0%; Score 821; DB 23; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.1e-86;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 6
AA06608
ID AAR06608 standard; protein; 160 AA.

AC AAR06608;
XX
DT 09-JAN-1991 (first entry)
XX
DE Human Superoxide Dismutase (hSOD).
XX
KW human Superoxide Dismutase (hSOD); pSOD104.
XX
OS Pichia pastoris.

XX Key Location/Qualifiers
FH Misc-difference 157..0
FT /label="stop" codon
FT Misc-difference 158..0
FT /label="stop" codon

XX W09009434-A.

XX 23-AUG-1990.

XX 12-FEB-1990; 90WO-US00781.

XX 12-FEB-1990; 90WO-US00781.

XX 13-FEB-1989; 89US-0311517.

XX (SALK) SALK INST FOR BIOL STUD.

XX Craig WS, Holtz GC, Davis GR;

XX WPI; 1990-275127/36.

XX N-PSDB; AAQ05849.

PT Prodn. of superoxidizedismutase - by culturing Pichia pastoris
XX yeast cells that have been transformed to express enzyme.

XX Disclosure; Fig 1; 54pp; English.

CC The gene for hSOD is under control of a promoter from a P.pastoris
CC gene to allow expression in P.pastoris e.g. the methanol-regulated
CC AOX1 gene promoter. The transcription termination signals from this
CC gene can also be used. P.pastoris cells transformed with a plasmid
CC carrying such a hSOD expression vector e.g. pSOD104, can produce
CC high levels of SOD to treat inflammation or to reduce superfusion
CC damage. The polypeptides produced in this way are properly folded,
CC in authentic form and have the correct size and surface charge.

CC Sequence is translated past the first stop codon in the specifica-
CC tion.

SQ Sequence 160 AA;

Query Match 100.0%; Score 821; DB 11; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.2e-86;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 3 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 62
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 63 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 122
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 123 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 156

RESULT 7

AAW33617
ID AAW33617 standard; Protein; 429 AA.

AC AAW33617;

XX 25-JUN-1998 (first entry)

XX SOD/p31pol structural region.

XX p25gag gene; enhanced promoter; gene expression; cytomegalovirus;
KW HIV; AIDS.

XX Human immunodeficiency virus type 1.

XX Key Location/Qualifiers
FH Protein 1..154
FT /note="SOD region"
FT Peptide 155..162
FT /note="linker sequence"
FT Protein 163..429
FT /note="p31pol region"

XX US5686888-A.

XX 18-NOV-1997.

XX 10-AUG-1994; 94US-0288336.

XX 24-DEC-1987; 87US-0138894.

XX 31-OCT-1984; 84US-0667501.

XX 30-JAN-1985; 85US-0696534.

XX 06-SEP-1985; 85US-0773447.

XX 17-AUG-1992; 92US-0931191.

XX 28-JUN-1993; 93US-0083391.

XX 17-AUG-1993; 93US-0107377.

XX 10-AUG-1994; 94US-0288336.

XX (CHIR) CHIRON CORP.

XX Chapman BS, Dina D, Haigwood NL, Luciw PA, Rosenberg S;

XX Thayer RM;

XX WPI; 1998-007982/01.

XX N-PSDB; AAW04741.

XX Enhanced promoter for gene expression - comprising cytomegalovirus
XX immediate early promoter plus intron
XX Example 4; Fig 24A-B; 99pp; English.


```

XX WPI; 1998-412999/35.
DR N-PSDB; AAV32580.
XX
XX New hybrid protein of superoxide dismutase and tetanus toxin
PT fragment C - having increased uptake by neurons and retention of
PT enzymatic activity in these cells, for treating neurological
PT diseases associated with oxidative stress
XX
XX Claim 7; Columns 23-26; 23pp; English.
XX
XX The present sequence represents an enzymatically active human
CC copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl
CC terminus with the tetanus toxin fragment C (TTC) moiety. The TTC
CC moiety constitutes amino acid residues 865-1315 of the tetanus
CC holotoxin. The hybrid protein, referred as SOD:Tet451, is claimed
CC to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic
CC activity; (b) the TTC moiety selectively binds to nerve cells and
CC allows uptake of the hybrid protein into these cells; and (c) it
CC retains substantial SOD enzymatic activity following cellular uptake.
CC SOD:Tet451 is claimed to be useful for treating neurological disorders
CC associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion,
CC epilepsy, Parkinson's and Huntington's diseases.
XX
XX Sequence 618 AA;
SQ
Query Match 100.0%; Score 821; DB 19; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.4e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
RESULT 10
AAW34481
ID AAW34481 standard; Protein; 1021 AA.
XX
AC AAW34481;
XX
XX 16-MAR-1998 (first entry)
XX
XX HCV antigen combination pSOD/c200/core.
XX
XX PCR primer; amplify; HCV; hepatitis C virus; antigen combination; NS3;
KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW NS4.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..154
FT /note= "hsod fragment"
FT Misc-difference 155..159
FT /note= "linker"
FT Misc-difference 160..899
FT /note= "c200 (amino acids 1192-1931 of HCV polyprotein)"
FT Misc-difference 1..902
FT /note= "linker"
FT Misc-difference 903..1021
FT /note= "c22 (amino acids 2-120 of HCV polyprotein)"
XX
XX US5683864-A.
PN

```

```

XX 04-NOV-1997.
PD XX
XX 07-JUL-1992; 92US-0910760.
PF XX
XX 07-JUL-1992; 92US-0910760.
PR XX
PR 18-NOV-1987; 87US-0122714.
PR 30-DEC-1987; 87US-0139886.
PR 26-FEB-1988; 88US-0161072.
PR 06-MAY-1988; 88US-0191263.
PR 26-OCT-1988; 88US-0263584.
PR 14-NOV-1988; 88US-0271450.
PR 17-MAR-1989; 89US-0325338.
PR 20-APR-1989; 89US-0341334.
PR 21-APR-1989; 89US-0353896.
PR 04-APR-1990; 90US-0504352.
XX
XX (CHIR ) CHIRON CORP.
PA
XX Choo Q, Houghton M, Kuo G;
PI
XX WPI; 1997-548976/50.
XX N-PSDB; AAT99982.
DR
XX Combination of three hepatitis C virus antigens - used for detection
PT of specific antibodies to diagnose infection
PT
XX Example 6; Column 59-68; 57pp; English.
XX
XX This sequence represents a Hepatitis C virus (HCV) antigen combination of
CC the invention. The HCV antigen combination comprises an antigen (Ag1)
CC comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV
CC polyprotein), or its immunologically reactive fragment containing at
CC least 8 aa. It also comprises two additional antigens from two different
CC polyprotein domains, including at least 8 aa from the NS3, NS4, S or NS5
CC domains of the polyprotein, corresponding, respectively, to aa 1050-1640;
CC 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively,
CC Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV
CC polyprotein. These antigen combinations are used diagnostically to detect
CC anti-HCV antibodies, using any standard immunoassay format. These antigen
CC combinations have a broader range of reactivity with antibodies than any
CC antigen individually.
XX
XX Sequence 1021 AA;
SQ
Query Match 100.0%; Score 821; DB 18; Length 1021;
Best Local Similarity 100.0%; Pred. No. 2.9e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
RESULT 11
AAW40039
ID AAW40039 standard; Protein; 1021 AA.
XX
AC AAW40039;
XX
XX 26-MAY-1998 (first entry)
DT
XX Fusion protein c200/c22.
DE
XX Hepatitis C virus C domain; HCV; immunological activity; c200/c22;
KW

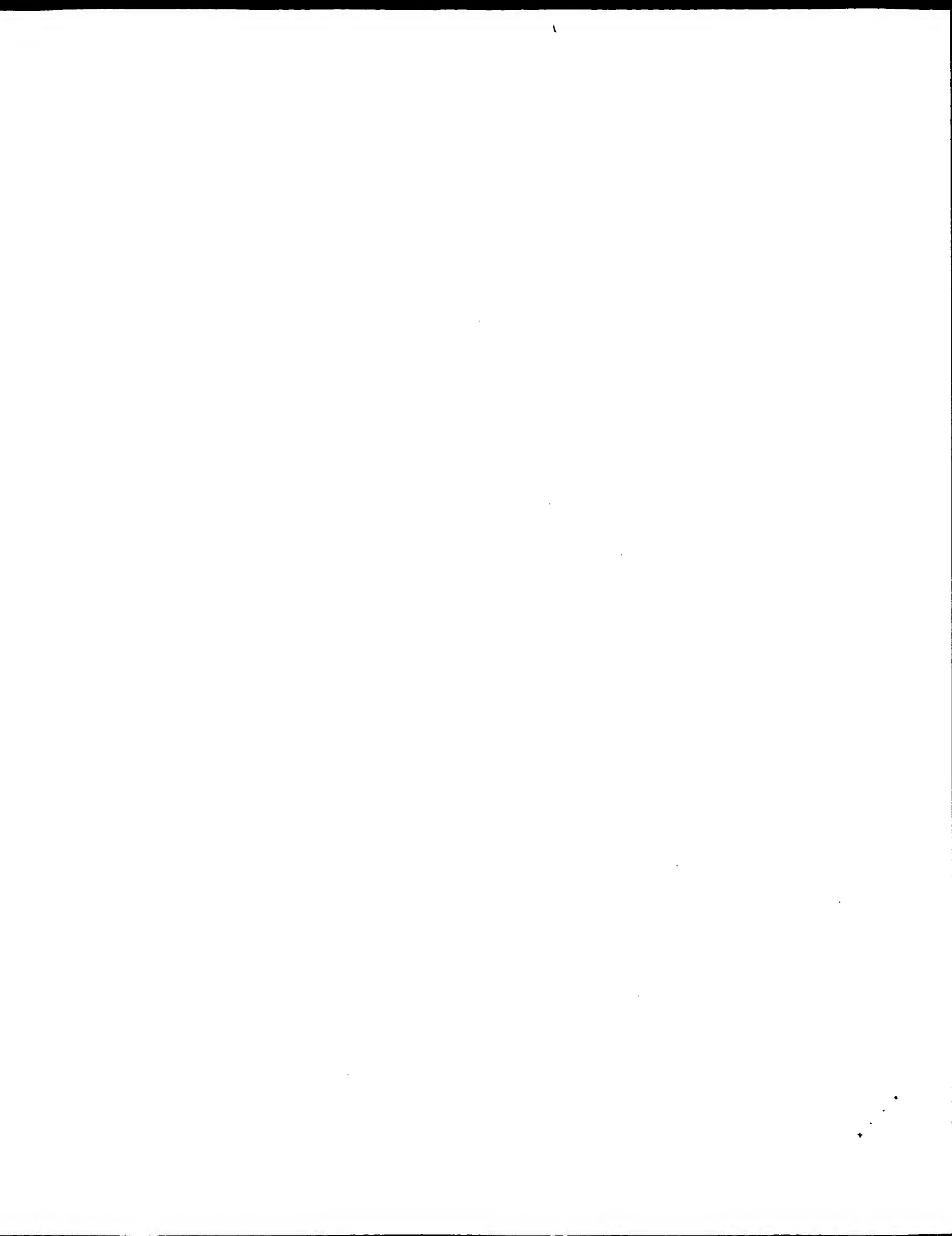
```


KW NS3 domain; NS4 domain; S domain; NS5 domain; fusion protein.
 XX Synthetic.
 OS Hepatitis virus.
 XX
 PN US5712087-A.
 XX
 XX 27-JAN-1998.
 XX
 XX 12-MAY-1995; 95US-0440519.
 XX
 XX 07-JUL-1992; 92US-0910760.
 PR 04-APR-1990; 90US-0504352.
 PR 12-MAY-1995; 95US-0440519.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Choo Q, Houghton M, Kuo G;
 PI WPI; 1998-119973/11.
 DR N-PSDB; AAV09990.
 XX
 PT Immunoassays for hepatitis C virus antibodies - using combinations
 PT of antigenic fragments of HCV polyprotein
 XX
 XX Example 6; Fig 4; 59pp; English.
 XX
 CC This sequence represents a fusion protein constructed from the hepatitis
 CC C virus core domain (which is situated at the carboxy terminus of the
 CC fusion protein) and a c200 construct (a fusion of the NS3 and NS5
 CC domains). This protein used in the construction of novel combinations of
 CC HCV antigens that have a broader range of immunological activity than
 CC any single HCV antigen. An example of such an antigen given in this
 CC specification comprises a first antigen containing at least 8 amino acids
 CC of the C domain of the HCV polyprotein and a second antigen comprising
 CC at least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or
 CC the NS5 domain of the HCV polyprotein in the form of a fusion protein, a
 CC physical mixture or bound to a solid matrix.
 XX
 SQ Sequence 1021 AA;
 Query Match 100.0%; Score 821; DB 19; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 2.9e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 RESULT 12
 AAE22050
 ID AAE22050 standard; Protein; 1021 AA.
 XX
 AC AAE22050;
 XX
 XX 16-JUL-2002 (first entry)
 XX
 XX pSOD/c200/core expression plasmid protein.
 DE
 XX Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
 KW NS4 domain; S domain; NS5 domain; pSOD/c200/core plasmid.
 XX
 XX Chimeric - Hepatitis C virus.
 OS Chimeric - Unidentified.

XX Key Location/Qualifiers
 FH Region 1..154
 FT /note= "hsOD"
 FT 155..159
 FT /note= "Linker region"
 FT 160..899
 FT /note= "HCV c200"
 FT 900..902
 FT /note= "Linker region"
 FT 903..1021
 FT /note= "HCV c22"
 XX
 PN US6312889-B1.
 XX
 XX 06-NOV-2001.
 XX
 XX 12-MAY-1995; 95US-0440549.
 XX
 XX 07-JUL-1992; 92US-0910760.
 PR 04-APR-1990; 90US-0504352.
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Houghton M, Choo Q, Kuo G;
 PI WPI; 2002-040268/05.
 DR N-PSDB; AAD35044.
 XX
 CC Combination of hepatitis C viral (HCV) antigens, useful in improved
 CC immunoassay for detecting HCV antibodies -
 XX Example 6; Fig 4; 58pp; English.
 XX
 CC The invention relates to combination of hepatitis C viral (HCV) antigens
 CC that have a broader range of immunological reactivity than any single HCV
 CC antigen. The combinations consist of an antigen from the C domain of the
 CC HCV polyprotein, and at least one additional HCV antigen from either the
 CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
 CC the form of fusion protein, a simple physical mixture, or the individual
 CC antigens commonly bound to a solid matrix. The combinations of antigens
 CC provides broad range immunoassays for anti-HCV antibodies. The invention
 CC therefore provides a method for detecting antibodies to HCV in a mammal
 CC suspected of containing such antibodies. The present sequence is a
 CC protein encoded by pSOD/c200/core expression plasmid DNA containing HCV
 CC coding sequence.
 XX
 SQ Sequence 1021 AA;
 Query Match 100.0%; Score 821; DB 23; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 2.9e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCIIIGRTLTV 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
 RESULT 13
 AAU76378
 ID AAU76378 standard; Protein; 1099 AA.
 XX
 AC AAU76378;
 XX
 XX 08-MAY-2002 (first entry)

XX HCV multiple epitope fusion antigen (MEFA) 7.1 protein sequence.
 DE Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
 XX Immunocassay solid support; multiple epitope fusion antigen; MEFA;
 KW non-structural protein.
 XX Hepatitis C virus.
 OS Synthetic.
 XX WQ200196870-A2.
 PN 20-DEC-2001.
 XX 14-JUN-2001; 2001WO-0519156.
 XX 15-JUN-2000; 2000US-212082P.
 PR 02-APR-2001; 2001US-280811P.
 PR 02-APR-2001; 2001US-280867P.
 XX (CHIR) CHIRON CORP.
 PA Chien DY, Arcangel P, Tandeske L, George-nascimento C, Coit D;
 PI Medina-selby A;
 XX WPI; 2002-090228/12.
 DR N-PSDB; ABK15345.
 XX Immunocassay solid support, useful for detecting hepatitis C virus
 PT infection in biological sample, comprises HCV NS3/4a conformational
 PT epitope and multiple epitope fusion antigen bound to the support -
 XX Claim 5; Fig 5; 92pp; English.
 XX The present invention relates to a new immunoassay solid support.
 CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
 CC conformational epitope and a multiple epitope fusion antigen (MEFA),
 CC bound to the support. The NS3/4a conformational epitope and/or
 CC MEFA reacts specifically with anti-HCV antibodies present in a biological
 CC sample from an HCV-infected individual. The immunoassay of the invention
 CC is useful for detecting hepatitis C virus infection in a biological
 CC sample. The method of the invention provides a sensitive, accurate
 CC diagnostic and prognostic tool to provide adequate patient care and to
 CC prevent transmission of HCV by blood and by blood products, or by
 CC personal contact. Use of NS3/4a conformational epitope in combination
 CC with MEFA, provides a sensitive and reliable method for detecting early
 CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
 CC masking problems, improving sensitivity in detecting antibodies by
 CC allowing a greater number of epitopes on a unit surface area of
 CC substrate, and improving substrate. Detection accuracy is increased and
 CC the incidence of false results is reduced because of the identification
 CC and the use of highly immunogenic HCV antigens which are present during
 CC the early stages of HCV seroconversion. The present amino acid sequence
 CC represents the multiple epitope fusion antigen (MEFA) 7.1 of the
 CC invention.
 XX Sequence 1099 AA;
 SQ Query Match 100.0%; Score 821; DB 23; Length 1099;
 Best Local Similarity 100.0%; Pred. No. 3.2e-85;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWVSGIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWVSGIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 QY 121⁶ HEKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154

RESULT 14
 AAP81019
 ID AAP81019 standard; protein; 154 AA.
 XX
 AC AAP81019;
 XX
 DT 01-NOV-1990 (first entry)
 XX
 DE Sequence of human superoxidizedismutase (hsOD) encoded by cDNA.
 XX
 KW Human superoxidizedismutase.
 XX
 FH Key Location/Qualifiers
 FT Region 2..154
 FT /note="Claimed in claim 3"
 XX
 PN JP63237790-A.
 XX
 XX 04-OCT-1988.
 XX 27-MAR-1987; 87JP-0073180.
 XX 27-MAR-1987; 87JP-0073180.
 XX (TOXN) TOYO JOZO KK.
 XX WPI; 1988-320034/45.
 DR N-PSDB; AAN81345.
 XX
 PT Prodn. of extracellular protein prod. -
 PT using human superoxidizedismutase linked to OMPA signal peptide
 PT and gram-negative bacteria
 XX
 XX ; Fig 3 Page 508; 23pp; Japanese.
 XX
 CC A gene which comprises the DNA encoding the signal peptide of the OMPA
 CC (AAN81344) linked upstream of DNA encoding an intracellular protein or
 CC peptide, esp. human superoxidizedismutase (hsOD) (AAN81345) is new. Prodn.
 CC of the intracellular protein/peptide comprises culturing bacteria contg.
 CC the new gene and accumulating the peptide/protein outside the cytoplasm
 CC of bacteria. Pref. the DNA having the basic sequence coding hSOD or its
 CC deriv. is obtd. by cloning cDNA of hSOD derived from human liver.
 CC Advantage is that purificn. of obtd. hSOD is easily carried out and the
 CC yield of the protein is higher than using conventional method.
 XX
 SQ Sequence 154 AA;
 Query Match 99.6%; Score 818; DB 9; Length 154;
 Best Local Similarity 99.4%; Pred. No. 4.6e-86;
 Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWVSGIKGLTEGLHGFHVEFGDNTAGCTS 60
 DB 1 MATKAVCVLKGDPVQGIINFEQKESGPKVWVSGIKGLTEGLHGFHVEFGDNTAGCTS 60
 QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLW 120
 QY 121 HEKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
 DB 121 HEKADDLGKGGNEESTKTGNAGSLRACGVIGIAQ 154
 RESULT 15
 AAP60235
 ID AAP60235 standard; Protein; 153 AA.
 XX
 AC AAP60235;
 XX
 DT 15-AUG-1991 (first entry)



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:19:14 ; Search time 2.01912 Seconds
(without alignments)
1539.035 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	10	US-09-904-987-6
2	821	100.0	1099	10	US-09-881-654-4
3	807	98.3	152	10	US-09-883-985-7
4	655.5	79.8	151	10	US-09-883-985-10
5	638	77.7	150	10	US-09-883-985-4
6	522	63.6	150	10	US-09-883-985-11
7	492	59.9	153	10	US-09-883-985-3
8	483	58.8	151	10	US-09-883-985-6
9	482	58.7	150	10	US-09-883-985-9
10	444.5	54.1	151	10	US-09-883-985-5
11	444.5	54.1	151	10	US-09-883-985-8
12	444	54.1	152	10	US-09-883-985-12
13	382.5	46.6	829	9	US-09-881-239-5
14	365	44.5	288	9	US-09-964-899-21
15	365	44.5	310	10	US-09-925-297-795
16	358	43.6	255	10	US-09-883-985-2
17	206.5	25.2	68	9	US-09-987-675-2
18	206.5	25.2	68	10	US-09-987-655-2
19	118	14.4	25	10	US-09-839-884-54

20	75.5	9.2	439	9	US-09-738-626-6639	Sequence 6639, Ap
21	75	9.1	14	9	US-09-826-290-262	Sequence 262, App
22	75	9.1	360	10	US-09-815-242-11254	Sequence 11254, A
23	72.5	8.8	1325	10	US-09-741-669-304	Sequence 304, App
24	72	8.8	261	10	US-09-940-037A-29	Sequence 29, Appl
25	72	8.8	552	9	US-09-961-721-2	Sequence 2, Appl
26	70.5	8.6	592	9	US-09-771-382-10	Sequence 10, Appl
27	70	8.5	579	9	US-10-108-605-215	Sequence 215, App
28	68.5	8.3	134	10	US-09-731-872-474	Sequence 474, App
29	68.5	8.3	442	9	US-10-117-417-6	Sequence 6, Appl
30	68.5	8.3	442	9	US-10-117-417-16	Sequence 16, Appl
31	68	8.3	170	12	US-10-109-885-3	Sequence 3, Appl
32	68	8.3	711	9	US-09-738-626-3507	Sequence 3507, Ap
33	67.5	8.2	595	10	US-09-864-761-38018	Sequence 38018, A
34	67.5	8.2	727	10	US-09-833-790-397	Sequence 397, App
35	67.5	8.2	1242	9	US-09-832-292-41	Sequence 41, Appl
36	67.5	8.2	26926	9	US-09-759-508B-2	Sequence 2, Appl
37	66.5	8.1	541	9	US-09-771-382-34	Sequence 34, Appl
38	66.5	8.1	592	9	US-09-771-382-2	Sequence 2, Appl
39	66.5	8.1	592	10	US-09-797-862-17	Sequence 17, Appl
40	66	8.0	135	10	US-09-815-242-10485	Sequence 10485, A
41	66	8.0	433	10	US-09-771-161A-229	Sequence 229, App
42	66	8.0	433	10	US-09-771-161A-230	Sequence 230, App
43	66	8.0	2568	10	US-09-866-108-3	Sequence 3, Appl
44	65.5	8.0	182	10	US-09-784-810A-29	Sequence 29, Appl
45	65.5	8.0	580	9	US-09-738-626-6264	Sequence 6264, Ap

ALIGNMENTS

RESULT 1

US-09-904-987-6

; Sequence 6, Application US/09904987

; Patent No. US20020037908A1

; GENERAL INFORMATION:

; APPLICANT: No. US20020037908A1actyl, Inc.

; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prop

; FILE REFERENCE: 42108/26146

; CURRENT APPLICATION NUMBER: US/09/904, 987

; CURRENT FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 154

; TYPE: PRT

; ORGANISM: homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441

; DATABASE ENTRY DATE: 2000-05-30

; RELEVANT RESIDUES: (1)..(154)

US-09-904-987-6

Query Match					100.0%; Score 821; DB 10; Length 154;
Best Local Similarity					100.0%; Pred. No. 4.3e-82;
Matches 154; Conservative					0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MATKAVCVLKGDPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVEFGDNTAGCTS	60		
Db	1	MATKAVCVLKGDPVQGIINFEQKESNPVKVWGSIKGLTEGLHGFHVEFGDNTAGCTS	60		
QY	61	AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGGDHCIIIGRTLTV	120		
Db	61	AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSTEDSVISLGGDHCIIIGRTLTV	120		
QY	121	HEKADDLGKGNESSTKTGNAGSLRACGVIGIAQ	154		
Db	121	HEKADDLGKGNESSTKTGNAGSLRACGVIGIAQ	154		

RESULT 2

US-09-881-654-4


```
;
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; STRANDEDNESS: single
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-883-985-10

Query Match          79.8%; Score 655.5; DB 10; Length 151;
Best Local Similarity 82.4%; Pred. No. 4e-64;
Matches 126; Conservative 7; Mismatches 15; Indels 5; Gaps 2;

QY 4 KAVCVLKGDPVQGIINFEQKESNGPVKVV--GSIKGLTEGLHGFHVEFGDNTAGCTSA 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 KAVCVLKGDPVQGTTHFEQKASGEP---WLSQITGLTEGQHGFGHVGQDNTQCTSA 58

QY 62 GPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVH 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 GPHFNPSKXKHGPKDEERHVGDLGNVTAGKGVANVSIEDRVISLGEHSIIIGRTMVVH 118

QY 122 EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 EKQDDLKGGNEESTKTGNAGSLACGVIGIIQ 151

RESULT 5
US-09-883-985-4
; Sequence 4, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
```

```
;
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-883-985-4

Query Match          77.7%; Score 638; DB 10; Length 150;
Best Local Similarity 82.0%; Pred. No. 3.2e-62;
Matches 123; Conservative 9; Mismatches 16; Indels 2; Gaps 1;

QY 3 TKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTSAG 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TKAVCVLKGDPVQGTTHFEAK--GDTVVVTGSIITGLTEGHDGFGHVGQDNTQCTSAG 58

QY 63 PHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGDHCHIGRTLVVHE 122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 PHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLGEYSIIGRTMVVHE 118

QY 123 KADDLGKGGNEESTKTGNAGSLACGVIGI 152
   || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 KPDDLGRGGNEESTKTGNAGSLACGVIGI 148

RESULT 6
US-09-883-985-11
; Sequence 11, Application US/09883985
; Patent No. US20020081288A1
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID
US-09-883-985-11

```

Query Match	63.6%	Score 522:	DB 10:	Length 150:
Best Local Similarity	66.9%	Pred. No. 1.3e-49:		
Matches	99:	Conservative	17:	Mismatches 30: Indels 2: Gaps 2:
QY	4	KACVVLKGGDPVQGIINFEQKESGPKVYKGSIKGLTEGLHGHHVFHGDNAGTACTSACP	63	
Db	2	KACVVLASGGDVKGVVRFEQD-DGDVTVEKTEGLDGNHGHIIHVFGDNTNGCLSAGP	60	
QY	64	HFNPLSRKHGPKPDERHHVGLGNVTADKCGVADVSTEDSVLSGDHCTIIGRTLVVHEK	123	
Db	61	HFNPQNKHGSPKADRHVGLGNVTAE-CGVQAFNFTDPOISLKGERSIIGRTAVVHEK	119	
QY	124	ADDLKGKGNBESTKKTGNAGSRKLAGCVTG	151	
Db	120	QDDLKGKGGDESLEKKTGNAGGRLAGCVTG	147	

RESULT.7

US-09-883-985-3
; Sequence 3, Application US/09883985
; Patent No. US20020081289A1
; GENERAL INFORMATION:

ROSEN, CRAIG A.
FRASER, CLAIRE M.
GOCAYNE, JEANNINE D.

SECRETARY GENERAL
TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600

ZIP: 20003
COMPUTER READABLE FORM:

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;

```

APPLICATION NUMBER: US/09/883,985
FILING DATE: 20-Jun-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/203,607
FILING DATE: 02-DEC-1998

FILING DATE: 02 DEC 1998
 APPLICATION NUMBER: US 08/722,050
 FILING DATE: 23-JAN-1997

FILING DATE: 23 JAN 1994
 APPLICATION NUMBER: US 08/225,757
 FILING DATE: 11-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.102

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600

```

:
: TELEFAX: (202) 371-2540
:
: INFORMATION FOR SEQ ID NO: 3:
:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 153 amino acids
:     TYPE: amino acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:

```

```

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-09-883-985-3

```

Query Match 59.9%; Score 492; DB 10; Length 153;
Best Local Similarity 60.0%; Pred. No. 2.4e-46;
Matches 90; Conservative 18; Mismatches 42; Indels

Qy	64	HFNPLSRKHGGPKDEERHVGLGNVTADKOGVADVSIEDSVISLSDGHCIIIGRTLWVHEK	123
		: : : : :	
Db	62	HENPTKQEHGAPESIRHVGLGNVVAGADGNAYNATDKLSUNGSHIIGRSMVIHEN	121

RESULT 8

US-09-883-985-6
; Sequence 6, Application US/09883985
; Patent No. US20020081288A1

GENERAL INFORMATION:
APPLICANT: YU, GUO-LIANG
ROSEN, CRAIG

FRASER, CLAIRE M.
GOCAYNE, JEANNINE D.

```

; SOURCE: HERRERA, E.
;
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
;
; NUMBER OF SEQUENCES: 16

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
;

CITY: WASHINGTON
STATE: DC

COUNTRY: USA
ZIP: 20005

```
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

;
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;

```

APPLICATION NUMBER: US/09/883,985
FILING DATE: 20-Jun-2001

```

;
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 09/203,607
FILING DATE: 02-DEC-1998

APPLICATION NUMBER: US 08/722,050
FILING DATE: 23-JAN-1997

APPLICATION NUMBER: US 08/225,757
FILING DATE: 11-APR-1994

[illegible]

QY 124 ADDLGKGGNEESTKTGNAGSRLACGVIGI 152
 |||||:| | |||| |::|||
 Db 121 PDDLKGGGHELSLSTGNAGGRVACGLIGI 149

RESULT 11

```

US-09-883-985-8
: Sequence 8, Application US/09883985
: Patent No. US200801288A1
: GENERAL INFORMATION:
: APPLICANT: YU, GUO-LIANG
: ROSEN, CRAIG A.
: FRASER, CLAIRE M.
: GOCAYNE, JEANNINE D.
: TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX P.L.L.C.
: STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER:
APPLICATION NUMBER: US/09/883,985
FILING DATE: 20-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/203,607
FILING DATE: 02-DEC-1998
APPLICATION NUMBER: US 08/722,050
FILING DATE: 23-JAN-1997
APPLICATION NUMBER: US 08/225,757
FILING DATE: 11-APR-1994

FILED DATE: 12 APR 1994
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36.688
REFERENCE/DOCKET NUMBER: 1488.1020003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600

TELEPHONE: (202) 371-2511
TELEFAX: (202) 371-2512
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 151 amino acids
TYPE: amino acid
STRANDEDNESS: single

SEQUENCE DESCRIPTION: SEQ ID NO: 8:
MOLECULE TYPE: protein
TOPOLOGY: linear

US-09-883-985-8

Query Match	54.1%;	Score 444.5;	DB 10;	Length 151;
Best Local Similarity	57.7%;	Pred. No. 3.4e-41;		
Matches	86;	Conservative 13;	Mismatches 49;	Indels 1;

	matches	86;	conservative	13;	mismatches	49;	indels	1;	gaps	
QV	4	KAVCVLKGDPVOGII	EOKESNGPVKVGSI	KGLTEGLHGFHV	HVEGDNTAGCT	SAGP	63			

db 2 KAVCVLNSSEGVSTGLFTQ-VGVAPTTVNGNISGLKPLGHGFHVHAGLDNTNCGMSTGP 60

Qy	64	HNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSTEDSVLSGDHCIGRTLVVHEK	123
Dh	61	HNYPAGKEHGAPEDEVHVGDLGNITVGEGTASFTTIDKOILPTGPOSTIGRAVVHAD	120

DB 8I HINFAGKEHGAPFEDVHRVGDLEGNII VGEDGIA
0Y 124 ADDLGKGGNEESTKTGNAGSRLACGVIGI 152

QY	124	ADDDGKGGNEESIKRIONAGSRRLACGVIGI	152
		: }	: : :
Db	121	PDDI,KGKGHELSKSTGNAGGRIACGIIGI	149

RESULT 12

```

US-09-883-985-12
; Sequence 12, Application US/09883985
; Patent No. US20020081288A1
;
; GENERAL INFORMATION:
;
; APPLICANT: YU, GUO-LIANG
;             ROSEN, CRAIG A.
;             FRASER, CLAIRE M.
;             GOCAYNE, JEANNINE D.
;
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
;
; NUMBER OF SEQUENCES: 16
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA

```

COUNTRY: USA
ZIP: 20005

ZIF: 20003
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/883,985
FILING DATE: 20-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/203,607

FILING DATE: 02-DEC-1998
APPLICATION NUMBER: US 08/722,050

FILING DATE: 23-JAN-1997
APPLICATION NUMBER: US 08/225,757

FILING DATE: 11-APR-1994
ATTORNEY/AGENT INFORMATION:

NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1020003
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

FORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-883-985-12

Q: same. No. 10. Length 152.
 54 19. 50000 111. 10 10. 152.

Query Match	54.1%	Score 444;	DB 10;	Length 152;
Best Local Similarity	56.4%	Pred. No. 3.9e-41;		
Matches	94.	Concentrations	17. Mismatch	48. Indels

Matches 84; Conservative 17; Mismatches 48; Indels

QY 4 KAVCVLKGDGPVQGIINFEQKESNGPVKWGSIRGLTEGLHGHHVFHEFGDNTAGG
 : ||| |||| |:: ||| : | | : | | : | | : | | : | | : | |
B4 2 GMAHVGSGVGVCKVKTGTCACEDDTMVCEPTKNCNDNAEEDTETUEEGCDATNCC
 : ||| |||| |:: ||| : | | : | | : | | : | | : | | : | |

Q: 54 HENDI SPWUCCDKDEBHMCDI CARITADKQCVADVISIEDSVISCDWCIIGPTI
Db 2 QAVAVLRGDAGVSGWKFEQASESEPTTVSYEIAGNSPAERFHIHEFGDATING

64 HFNP LSRKHGGPKDEERHVGDLGNVTADKDGADVSDSVLSGDHCILGR
QY
63 HFNDKPKKTHGADPDDEEVHVGDMCNKTDNCAVCKSEKDSIKIKIPDTSVWGRSV

DB 62 HFNPFRKTHGAPTDEVHRHVGDMGNVKTDENGVAKGSFKDSLKLIKLGPTSVVGRS
124 ADDICKCCNEESTKTCNAGSPICCCVIGI 152

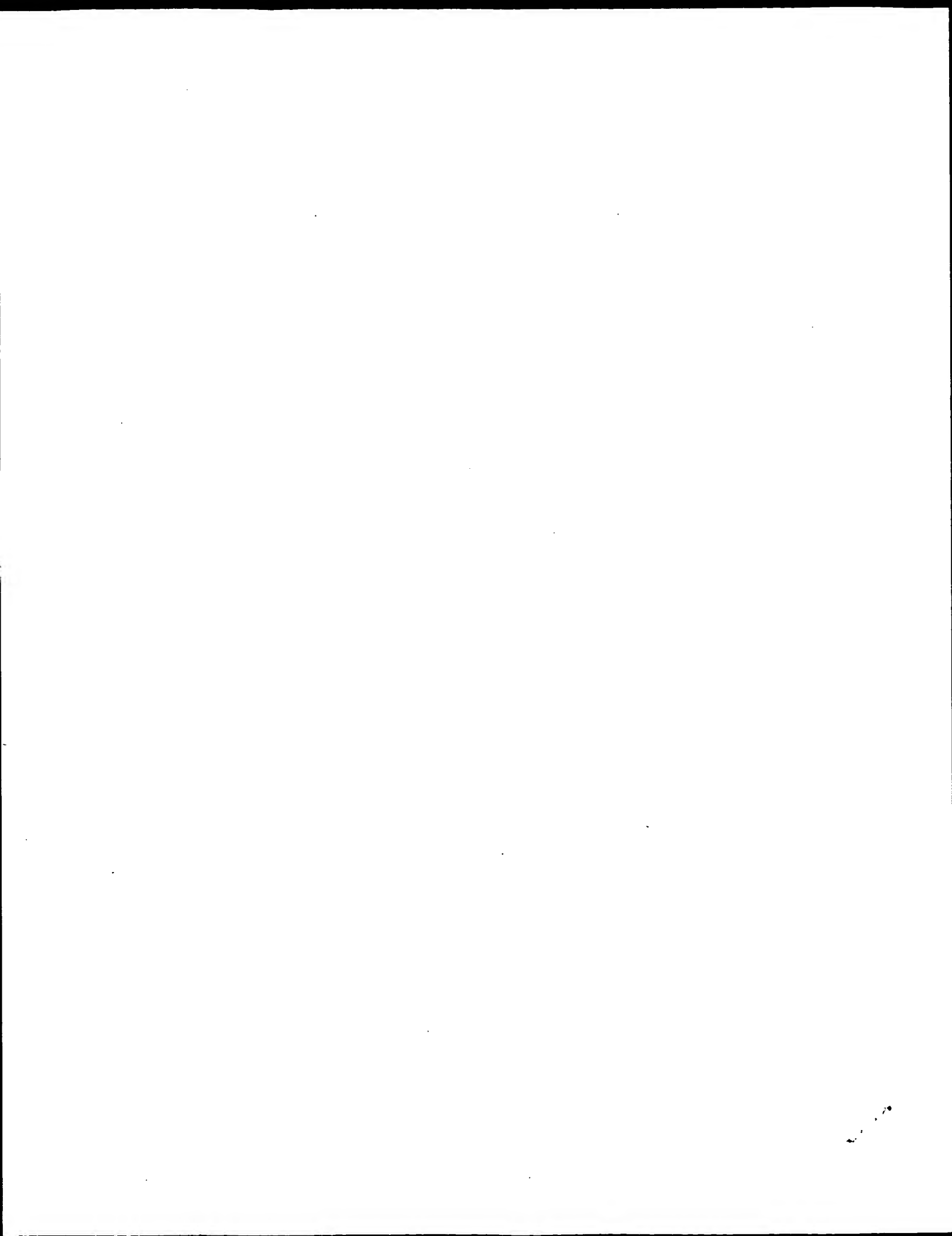
QY	124	ADDLGRGGNEESTKKTGNAGSR	LACGVIGI	152
Dh	122	ADDLGRGGTDEESIKKTGNAGSR	RRACGVIGI	150

DB 122 QDDLGLKGDTEESLKTGNAGPRPACGVIGI 150

RESULT 13
US-00-881-230-5

US-09-881-239-5
; Sequence 5, Application US/09881239
; Publication NO. US20020192639A1

; Publication No. US20020192639A1



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:14:29 ; Search time 39,9197 Seconds
(without alignments)
2487.216 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDPVQGIIN.....STKTGNAGSLRACGVIGIAQ 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main :*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	154	1	PCT-US01-14772-3
2	821	100.0	154	1	PCT-US02-10824-198
3	821	100.0	154	1	PCT-US02-22032-6
4	821	100.0	154	21	US-09-744-016A-2
5	821	100.0	154	21	US-09-744-016A-16
6	821	100.0	154	21	US-09-791-537-38416

7	821	100.0	154	23	US-09-904-987-6	Sequence 6, Appli
8	821	100.0	156	21	US-09-760-475-2306	Sequence 2306, Ap
9	821	100.0	668	13	US-08-993-010-2	Sequence 2, Appli
10	821	100.0	839	21	US-09-791-537-91118	Sequence 91118, A
11	821	100.0	1099	22	US-09-881-654-4	Sequence 4, Appli
12	816	99.4	153	3	US-07-768-912-13	Sequence 13, Appli
13	816	99.4	153	21	US-09-744-016A-14	Sequence 14, Appli
14	816	99.4	153	24	US-10-018-770-1	Sequence 1, Appli
15	816	99.4	154	21	US-09-791-537-151133	Sequence 151133,
16	814	99.1	154	21	US-09-791-537-118562	Sequence 118562,
17	814	99.1	154	27	US-60-389-987-185	Sequence 185, App
18	814	99.1	154	27	US-60-412-418-185	Sequence 185, App
19	812	98.9	153	21	US-09-791-537-99028	Sequence 99028, A
20	808	98.4	153	21	US-09-791-537-93750	Sequence 93750, A
21	807	98.3	152	16	US-09-203-607-7	Sequence 7, Appli
22	807	98.3	152	22	US-09-883-985-7	Sequence 7, Appli
23	807	98.3	179	21	US-09-791-537-68361	Sequence 68361, A
24	802	97.7	154	21	US-09-791-537-53587	Sequence 53587, A
25	797	97.1	153	21	US-09-791-537-83751	Sequence 83751, A
26	797	97.1	154	21	US-09-791-537-152870	Sequence 152870,
27	797	97.1	171	21	US-09-791-537-50085	Sequence 50085, A
28	797	97.1	841	3	US-07-680-296-86	Sequence 86, Appli
29	797	97.1	841	22	US-09-884-455-86	Sequence 86, Appli
30	797	97.1	841	22	US-09-884-456-86	Sequence 86, Appli
31	793	96.6	153	21	US-09-791-537-38270	Sequence 38270, A
32	790	96.2	153	27	US-60-389-987-74	Sequence 74, Appli
33	790	96.2	153	27	US-60-412-418-74	Sequence 280, App
34	788	96.0	200	1	PCT-US01-14826-280	Sequence 13968, A
35	777	94.6	153	21	US-09-791-537-13968	Sequence 13968, A
36	773	94.2	153	21	US-09-791-537-151958	Sequence 151958,
37	763	92.9	153	21	US-09-791-537-150043	Sequence 150043,
38	744	90.6	154	4	US-08-023-980-38	Sequence 38, Appli
39	744	90.6	154	6	US-08-204-052-51	Sequence 51, Appli
40	689.5	84.0	153	8	US-08-486-953-51	Sequence 51, Appli
41	684.5	83.4	152	21	US-09-791-537-42081	Sequence 42081, A
42	684.5	83.4	152	21	US-09-791-537-42336	Sequence 42336, A
43	684.5	83.4	152	21	US-09-791-537-85330	Sequence 85330, A
44	680	82.8	154	21	US-09-791-537-43187	Sequence 43187, A
45	678	82.6	154	21	US-09-791-537-43206	Sequence 43206, A

ALIGNMENTS

RESULT 1

PCT-US01-14772-3
; Sequence 3, Application PC/TUS0114772
; GENERAL INFORMATION:
; APPLICANT: Genesuisse Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y.
; APPLICANT: Bentivegna, Steven C.
; APPLICANT: Kliem, Stefanie E.
; APPLICANT: Koshiy, Beena
; APPLICANT: Parks, Katie E.
; TITLE OF INVENTION: HAPLOTYPES OF THE SODI GENE
; FILE REFERENCE: MMH-0429PCT SODI
; CURRENT APPLICATION NUMBER: PCT/US01/14772
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 60/202,491
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-14772-3

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQESNCPVKWGSIKLTGLHGFHVFEGDNTAGCTS 60

|||||
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 2
PCT-US02-10824-198
; Sequence 198, Application PC/TUS0210824
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Prostate Cancer Expression Profiles
; FILE REFERENCE: 90 206 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/10824
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/281,732
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/281,731
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 198
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10824-198

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 3
PCT-US02-22032-6
; Sequence 6, Application PC/TUS0222032
; GENERAL INFORMATION:
; APPLICANT: Novactyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Protein Assembly or Aggr
; FILE REFERENCE: 42108/34520
; CURRENT APPLICATION NUMBER: PCT/US02/22032
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 09/904,987
; PRIOR FILING DATE: 12 July 2001
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 154
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441
; DATABASE ENTRY DATE: 2000-05-30
; RELEVANT RESIDUES: (1)..(154)
PCT-US02-22032-6

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4
US-09-744-016A-2
; Sequence 2, Application US/09744016A
; GENERAL INFORMATION:
; APPLICANT: Dr. Voelkel, Helge
; TITLE OF INVENTION: Method for screening of modulators of calcineurin
; TITLE OF INVENTION: activity
; FILE REFERENCE: A34157PCT
; CURRENT APPLICATION NUMBER: US/09/744,016A
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: EP98113876
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-744-016A-2

Query Match 100.0%; Score 821; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 5
US-09-744-016A-16
; Sequence 16, Application US/09744016A
; GENERAL INFORMATION:
; APPLICANT: Dr. Voelkel, Helge
; TITLE OF INVENTION: Method for screening of modulators of calcineurin
; TITLE OF INVENTION: activity
; FILE REFERENCE: A34157PCT
; CURRENT APPLICATION NUMBER: US/09/744,016A
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: EP98113876
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-744-016A-16


```

Query Match          100.0%; Score 821; DB 21; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.2e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

	Qy	1	MATKACVVLKGDGPVGIIINFEOKESNGPVKWGSKTGLTEGLHGPHVFHEFGDNTAGCTS	60
	Dd	1	MATKACVVLKGDGPVGIIINFEOKESNGPVKWGSKTGLTEGLHGPHVFHEFGDNTAGCTS	60
	Qy	61	AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGADVSIEDSVISLSDGHCIIGRTLTV	120
	Dd	61	AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGADVSIEDSVISLSDGHCIIGRTLTV	120
	Qy	121	HKADDLKGKNPEESTKTGNASRLACGVIGIAQ	154
	Dd	121	HKADDLKGKNPEESTKTGNASRLACGVIGIAQ	154

```

RESULT 6
US-09-791-537-38416
; Sequence 38416, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38416
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-38416

```

[illegible]

```

RESULT 7
US-09-904-987-6
; Sequence 6, Application US/09904987
; GENERAL INFORMATION:
; APPLICANT: Novacetyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepath
; TITLE OF INVENTION: Protein Assembly or Aggregation
; FILE REFERENCE: 42108/26146
; CURRENT APPLICATION NUMBER: US/09/904, 987
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 154
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / P00441

```

```

; DATABASE ENTRY DATE: 2000-05-30
; ; RELEVANT RESIDUES: (1)..(154)
US-09-904-987-6

```

Query Match	100.0%;	Score 821;	DB 23;	Length 154;
Best Local Similarity	100.0%;	Pred. No. 2.2e-84;		
Matches 154;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps
Qy	1	MATKAVCVLKGDCGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFPHVHEFGDNTAGCTS	60	
Db	1	MATKAVCVLKGDCGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFPHVHEFGDNTAGCTS	60	
Qy	61	AGPHFNPLSRKHGCPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV	120	
Db	61	AGPHFNPLSRKHGCPKDEERHVGDLGNVTADKGVADVSTEDSVISLSDGHCIIIGRTLTV	120	
Qy	121	HEKADDLKGKGNBESTKTGNAGSRLACGVIGIAQ	154	
Db	121	HEKADDLKGKGNBESTKTGNAGSRLACGVIGIAQ	154	

```

RESULT 8
US-09-760-475-2306
; Sequence 2306, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: Ptz49
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2306
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2306

```

	Query Match	100.0%;	Score 821;	DB 21;	Length 156;
	Best Local Similarity	100.0%;	Pred. No. 2.2e-84;		
	Matches 154;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MATKAVCVLKGDPVQGIIINEFOEKESGPNVKWGSIKGLTEGLGHFHVHFGDNTAGCTS	60		
Dd	3	MATKAVCVLKGDPVQGLIINEFOEKESGPNVKWGSIKGLTEGLGHFHVHFGDNTAGCTS	62		
Qy	61	AGPHFNPLSRKHGGPKDEERHWGDLGNTADKDGADVYSIEDSVISLSDGHCIIIGRTLIV	120		
Dd	63	AGPHFNPLSRKHGGPKDEERHWGDLGNTADKDGADVYSIEDSVISLSDGHCIIIGRTLIV	122		
Qy	121	HEKADDLKGGNEESTTKTGNAGSLACCGVIGIAQ	154		
Dd	123	HEKADDLKGGNEESTTKTGNAGSLACCGVIGIAQ	156		

```

RESULT 9
US-08-993-010-2
: Sequence 2, Application US/08993010
: GENERAL INFORMATION:
: APPLICANT: Quan, Stella
: APPLICANT: Valenzuela, Pablo
: APPLICANT: Polito, Alan
: TITLE OF INVENTION: HELICOBACTER PYLORI DIAGNOSTICS
: FILE REFERENCE: 1360.002
: CURRENT APPLICATION NUMBER: US/08/993,010
: CURRENT FILING DATE: 1997-12-18
: EARLIER APPLICATION NUMBER: 60/033,707
: EARLIER FILING DATE: 1996-12-19
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 2
: LENGTH: 668

```



```
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-993-010-2

Query Match      100.0%; Score 821; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.6e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 10
US-09-791-537-91118
; Sequence 91118, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91118
; LENGTH: 839
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-791-537-91118

Query Match      100.0%; Score 821; DB 21; Length 839;
Best Local Similarity 100.0%; Pred. No. 2.3e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 11
US-09-881-654-4
; Sequence 4, Application US/09881654
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANGEL, Philip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIEMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PPI7039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
```

```
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1099
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
US-09-881-654-4

Query Match      100.0%; Score 821; DB 22; Length 1099;
Best Local Similarity 100.0%; Pred. No. 3.3e-83;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||||
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60

QY 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120
   |||||||
Db 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGDVADVSIEDSVISLSDHCHCIIGRTLTV 120

QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||||
Db 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 12
US-07-768-912-13
; Sequence 13, Application US/07768912
; GENERAL INFORMATION:
; APPLICANT: Smith, Geoffrey
; TITLE OF INVENTION: VACCINIA VECTORS, VACCINIA GENES AND
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/768,912
; APPLICATION NUMBER: US/07/768,912
; FILING DATE: 19911112
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-55557
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-768-912-13

Query Match      99.4%; Score 816; DB 3; Length 153;
```


Best Local Similarity 100.0%; Pred. No. 7.9e-84;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 153

RESULT 13

US-09-744-016A-14

; Sequence 14, Application US/09744016A

; GENERAL INFORMATION:

; APPLICANT: Dr. Voelkel, Helge

; TITLE OF INVENTION: Method for screening of modulators of calcineurin

; FILE REFERENCE: A34157PCT

; CURRENT APPLICATION NUMBER: US/09/744,016A

; PRIOR FILING DATE: 2001-10-03

; PRIOR FILING DATE: 1998-07-22

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-744-016A-14

Query Match 99.4%; Score 816; DB 21; Length 153;

Best Local Similarity 100.0%; Pred. No. 7.9e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 153

RESULT 14

US-10-018-770-1

; Sequence 1, Application US/10018770

; GENERAL INFORMATION:

; APPLICANT: LIT Institute Co., Ltd., and Selkagaku Corporation

; TITLE OF INVENTION: Pharmaceutical composition containing lecithinized-superoxide dis

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/018,770

; CURRENT FILING DATE: 2001-12-17

; NUMBER OF SEQ ID NOS: 1

; SEQ ID NO 1

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-018-770-1

Query Match

Best Local Similarity 99.4%; Score 816; DB 24; Length 153;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
DB 1 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 60
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 121
DB 61 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 120
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 153

RESULT 15

US-09-791-537-151133

; Sequence 151133, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 151133

; LENGTH: 154

; TYPE: PRT

; ORGANISM: pdb 1SPDA

; NAME/KEY: misc_feature

; LOCATION: (1)..(1)

; OTHER INFORMATION: X is an unknown amino acid

US-09-791-537-151133

Query Match 99.4%; Score 816; DB 21; Length 154;

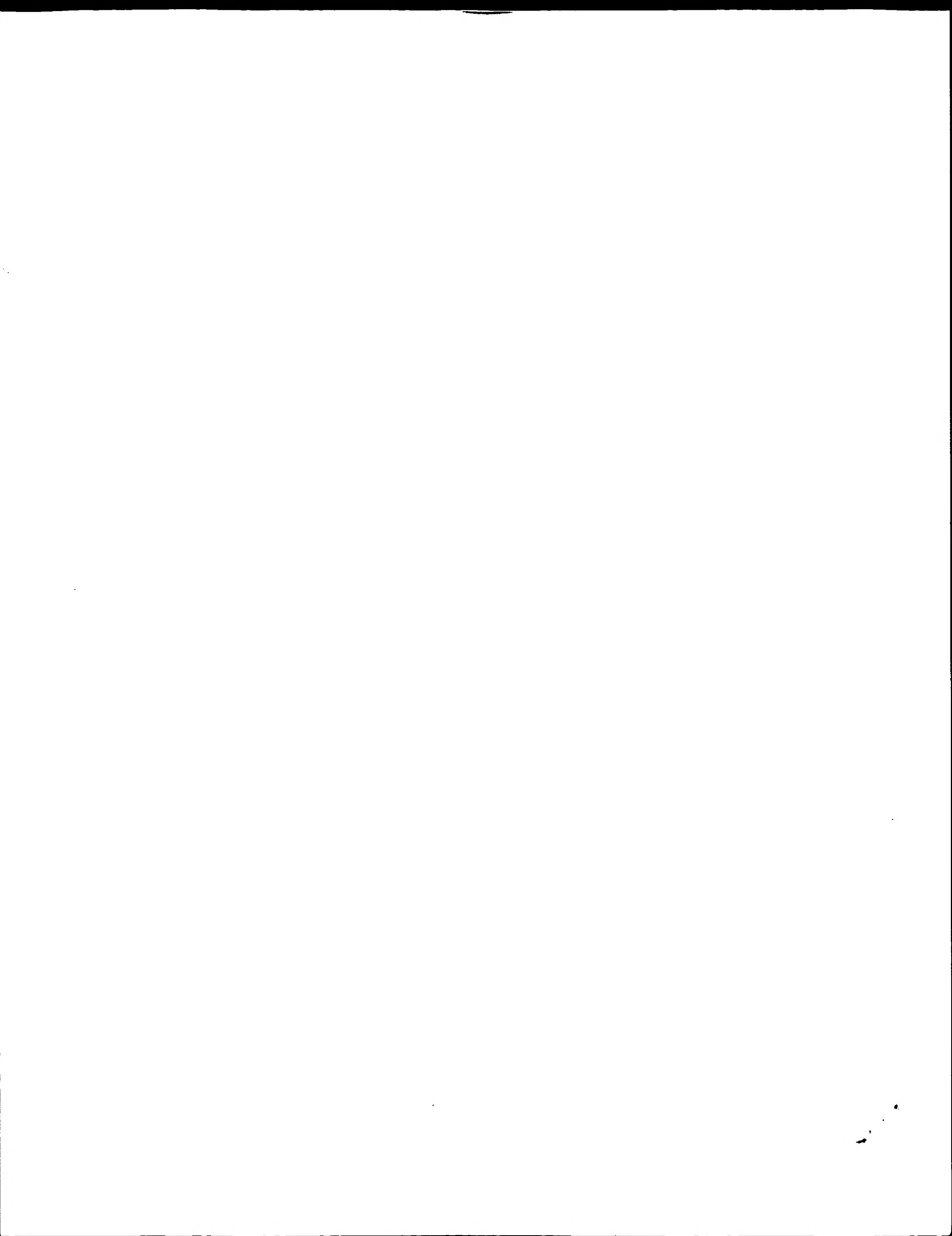
Best Local Similarity 100.0%; Pred. No. 8e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
DB 2 ATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHEFGDNTAGCTSA 61
QY 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 121
DB 62 GPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSDIEDSVISLSDGHCIIIGRTLTVVH 121
QY 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 122 EKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

Search completed: January 28, 2003, 14:37:27

Job time : 41.9197 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:16:13 ; Search time 4.50096 Seconds
(without alignments)
2701.343 Million cell updates/sec

Title: US-09-904-987-6
Perfect score: 821
Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKTGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 394613
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pap:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pap:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pap:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pap:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pap:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pap:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	821	100.0	154	1	PCT-US02-35563-20	Sequence 20, Appl
2	821	100.0	236	1	PCT-US02-33121-46	Sequence 46, Appl
3	821	100.0	236	6	US-10-272-459-46	Sequence 46, Appl
4	821	100.0	352	1	PCT-US02-33121-45	Sequence 45, Appl
5	821	100.0	352	6	US-10-272-459-45	Sequence 45, Appl
6	821	100.0	382	1	PCT-US02-33121-47	Sequence 47, Appl
7	821	100.0	382	6	US-10-272-459-47	Sequence 47, Appl
8	821	100.0	652	1	PCT-US02-33121-48	Sequence 48, Appl
9	821	100.0	652	6	US-10-272-459-48	Sequence 48, Appl
10	682	83.1	127	5	US-09-513-999C-5289	Sequence 5289, Ap
11	644	78.4	128	5	US-09-724-676-81086	Sequence 81086, A
12	644	78.4	128	5	US-09-724-676-81097	Sequence 81097, A
13	644	78.4	128	5	US-09-724-676-81107	Sequence 81107, A
14	644	78.4	128	5	US-09-724-676-81112	Sequence 81112, A
15	644	78.4	128	5	US-09-724-676-81111	Sequence 81111, A
16	644	78.4	128	5	US-09-724-676A-81107	Sequence 81107, A
17	528.5	64.4	150	5	US-09-724-676A-81107	Sequence 81107, A
18	528.5	64.4	150	5	US-09-724-676A-81112	Sequence 81112, A
19	449.5	54.8	88	5	US-09-724-676-81111	Sequence 81111, A
20	449.5	54.8	88	5	US-09-724-676A-81111	Sequence 81111, A
21	447.5	54.5	88	5	US-09-724-676-81089	Sequence 81089, A
22	447.5	54.5	88	5	US-09-724-676-81100	Sequence 81100, A
23	447.5	54.5	88	5	US-09-724-676A-81089	Sequence 81089, A
24	447.5	54.5	88	5	US-09-724-676A-81100	Sequence 81100, A
25	442	53.8	80	5	US-09-724-676-81087	Sequence 81087, A
26	442	53.8	80	5	US-09-724-676-81088	Sequence 81088, A

27	442	53.8	80	5	US-09-724-676-81098	Sequence 81098, A
28	442	53.8	80	5	US-09-724-676-81099	Sequence 81099, A
29	442	53.8	80	5	US-09-724-676-81108	Sequence 81108, A
30	442	53.8	80	5	US-09-724-676-81109	Sequence 81109, A
31	442	53.8	80	5	US-09-724-676A-81087	Sequence 81087, A
32	442	53.8	80	5	US-09-724-676A-81088	Sequence 81088, A
33	442	53.8	80	5	US-09-724-676A-81098	Sequence 81098, A
34	442	53.8	80	5	US-09-724-676A-81099	Sequence 81099, A
35	442	53.8	80	5	US-09-724-676A-81108	Sequence 81108, A
36	442	53.8	80	5	US-09-724-676A-81109	Sequence 81109, A
37	365	44.5	218	5	US-09-724-676-78095	Sequence 78095, A
38	365	44.5	218	5	US-09-724-676A-78095	Sequence 78095, A
39	305.5	37.2	80	5	US-09-724-676-81101	Sequence 81101, A
40	305.5	37.2	80	5	US-09-724-676A-81101	Sequence 81101, A
41	303	36.9	66	5	US-09-724-676-81096	Sequence 81096, A
42	303	36.9	66	5	US-09-724-676A-81096	Sequence 81096, A
43	303	36.9	69	5	US-09-724-676-81090	Sequence 81090, A
44	303	36.9	69	5	US-09-724-676-81110	Sequence 81110, A
45	303	36.9	69	5	US-09-724-676A-81090	Sequence 81090, A

ALIGNMENTS

RESULT 1

PCT-US02-35563-20
; Sequence 20, Application PC/TUS0235563
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc
; TITLE OF INVENTION: REGULATED PROSTATE CANCER GENES
; FILE REFERENCE: Oct 90 803 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/35563
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/331,042
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/331,041
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/340,251
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/344,791
; PRIOR FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 20
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-35563-20

Query Match 100.0%; Score 821; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 3.8e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATKAVCVLKGDGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS	60
Db	1	MATKAVCVLKGDGPVQGIINFEOKESNGPVKVMGSIKGLTEGLHGFHVEFGDNTAGCTS	60
QY	61	AGPHFNPLSRKKGPKDEERHVGDLGNVTADKDGADVSDIEDSVISLSDGHCIIIGRTLVV	120
Db	61	AGPHFNPLSRKKGPKDEERHVGDLGNVTADKDGADVSDIEDSVISLSDGHCIIIGRTLVV	120
QY	121	HEKADDLGKGNNESTKTGNAGSLACGVIGIAQ	154
Db	121	HEKADDLGKGNNESTKTGNAGSLACGVIGIAQ	154

RESULT 2

PCT-US02-33121-46
; Sequence 46, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF


```
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; TYPE: PRT
; LENGTH: 236
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 24.9 kDa
PCT-US02-33121-46

Query Match          100.0%; Score 821; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.3e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||

RESULT 3
US-10-272-459-46
; Sequence 46, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 24.9 kDa
US-10-272-459-46

Query Match          100.0%; Score 821; DB 6; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.3e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||

RESULT 4
PCT-US02-33121-45
; Sequence 45, Application PC/TUS0233121
; GENERAL INFORMATION:
```

```
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
PCT-US02-33121-45

Query Match          100.0%; Score 821; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.9e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||

RESULT 5
US-10-272-459-45
; Sequence 45, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match          100.0%; Score 821; DB 6; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.9e-61;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFEFGDNTAGCTS 60
   |||||
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
   |||||
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
   |||||

RESULT 6
```



```

Db      61  AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHIIGRTLVV 120
Qy      121  HEKADDLGRKGNEESTKTKGNAGSRLACGVIGIAQ 154
          |||||
Db      121  HEKADDLGRKGNEESTKTKGNAGSRLACGVIGIAQ 154
          |||||

RESULT 10
US-09-513-999C-5289
; Sequence 5289, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5289
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5289

Query Match      83.1%; Score 682; DB 5; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.1e-49;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGRFVHFEFGDNTAGCTS 60
          |||||
Db      1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGRFVHFEFGDNTAGCTS 60
          |||||

Qy      61  AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHIIGRTLVV 120
          |||||
Db      61  AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHIIGRTLVV 120
          |||||

Qy      121  HEKADDL 127
          |||||
Db      121  HEKADDL 127
          |||||

RESULT 11
US-09-724-676-81086
; Sequence 81086, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81086
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-81086

Query Match      78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGRFVHFEFGDNTAGCTS 60
          |||||
Db      1 MATKAVCVLKGDPVQGIINFEQKESGPKVWGSIKGLTEGLHGRFVHFEFGDNTAGCTS 60
          |||||

Qy      61  AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHIIGRTLVV 120
          |||||
Db      61  AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSGDHCHIIGRTLVV 120
          |||||

```


; SEQ ID NO 81086
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-81086

Query Match 78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGADVADSVIEDSVISLSGDHCCIIGRTLTV 120
Db 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGADVADSVIEDSVISLSGDHCCIIGRTLTV 120

RESULT 15

US-09-724-676A-81097
; Sequence 81097, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81097
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-81097

Query Match 78.4%; Score 644; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
Db 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGADVADSVIEDSVISLSGDHCCIIGRTLTV 120
Db 61 AGPHFNPLSRKHGKGPDEERHVDLGNVTADKDGADVADSVIEDSVISLSGDHCCIIGRTLTV 120

Search completed: January 28, 2003, 14:39:22
Job time : 6.50096 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:10:57 : Search time 3.19694 Seconds
(without alignments)
1417.333 Million cell updates/sec

Title: US-09-904-987-6

Perfect score: 821

Sequence: 1 MATKAVCVLKGDGPVQGIIN.....STKTGNAGSLACGVIGIAQ 154

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	821	100.0	154	4	US-09-126-109-4
2	821	100.0	618	1	US-08-668-381A-5
3	821	100.0	1021	1	US-07-910-760-12
4	821	100.0	1021	1	US-08-440-519-12
5	821	100.0	1021	4	US-08-440-549-12
6	816	99.4	153	4	US-08-679-493A-204
7	816	99.4	153	6	5290690-6
8	812	98.9	154	6	5290690-5
9	807	98.3	152	2	US-08-722-050-7
10	801.5	97.6	152	6	5196335-1
11	797	97.1	841	1	US-08-350-884-86
12	797	97.1	841	1	US-08-709-173-86
13	797	97.1	841	2	US-08-709-177-86
14	772	94.0	144	6	5252476-5
15	758	92.3	152	6	5171680-3
16	744	90.6	154	2	US-08-023-980B-43
17	744	90.6	154	2	US-08-486-953A-51
18	684.5	83.4	152	4	US-08-679-493A-203
19	673	82.0	153	4	US-08-679-493A-202
20	670	81.6	153	4	US-08-679-493A-201
21	662	80.6	153	4	US-08-679-493A-207
22	655.5	79.8	151	2	US-08-722-050-10
23	650.5	79.2	152	4	US-08-679-493A-206
24	647	78.8	151	4	US-09-202-832-16
25	647	78.8	151	4	US-08-679-493A-205
26	638	77.7	150	2	US-08-722-050-4
27	638	77.7	151	6	5290690-7

28	589.5	71.8	166	4	US-08-679-493A-209
29	558.5	68.0	152	4	US-08-679-493A-208
30	527.5	64.3	149	4	US-08-679-493A-200
31	522	63.6	150	2	US-08-722-050-11
32	521	63.5	150	4	US-08-679-493A-199
33	492	59.9	153	2	US-08-722-050-3
34	483	58.8	151	2	US-08-722-050-9
35	482	58.7	150	2	US-08-722-050-9
36	475.5	57.9	151	4	US-08-679-493A-192
37	471.5	57.4	151	4	US-08-679-493A-210
38	469.5	57.2	151	4	US-08-679-493A-191
39	464.5	56.6	201	4	US-08-679-493A-190
40	453.5	56.3	202	4	US-08-679-493A-197
41	455.5	55.5	218	4	US-08-679-493A-195
42	453.5	55.2	217	4	US-08-679-493A-196
43	448	54.6	221	4	US-08-679-493A-198
44	444.5	54.1	151	2	US-08-722-050-5
45	444.5	54.1	151	2	US-08-722-050-8

ALIGNMENTS

RESULT 1

US-09-126-109-4

; Sequence 4, Application US/09126109

; Patent No. 6171856

; GENERAL INFORMATION:

; APPLICANT: Thigpen, Anice

; APPLICANT: Holmeier, Hans-Ewald

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Unger, Roger H.

; APPLICANT: Shimabukuro, Michio

; APPLICANT: Chen, Guaxun

; APPLICANT: Rhodes, Christopher J.

; APPLICANT: Hugl, Sigrun R.

; APPLICANT: Cousin, Sharon

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING

; TO NO-MEDIATED CYTOTOXICITY

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER:

; APPLICATION NUMBER: US/09/136,109

; FILING DATE: 30-JUL-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/055,092

; FILING DATE: 30-JUL-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US Unknown

; FILING DATE: 03-MAR-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: McMillian, Nabeela R.

; REGISTRATION NUMBER: P-43,363

; REFERENCE/DOCKET NUMBER: UTSD:560

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 154 amino acids

; TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
US-09-126-109-4

Query Match 100.0%; Score 821; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 7.7e-85;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 2

US-08-668-381A-5
Sequence 5, Application US/08668381A
Patent No. 5780024
GENERAL INFORMATION:

APPLICANT: Brown, Robert H.
APPLICANT: Fishman, Paul S.
APPLICANT: Francis, Jonathan W.
APPLICANT: Hosler, Betsy A.
TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA

ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/668,381A
FILING DATE: 21-JUN-1996
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/000,473
FILING DATE: 23-JUN-1995
ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,164
REFERENCE/DOCKET NUMBER: 00786/269001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 618 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-668-381A-5

Query Match 100.0%; Score 821; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 5e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60

DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 3

US-07-910-760-12
Sequence 12, Application US/07910760
Patent No. 5683864

GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.

ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/910,760
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-910-760-12

Query Match 100.0%; Score 821; DB 1; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
DB 1 MATKAVCVLKGDPVQGIINFEQKESNGPVKVGSIKGLTEGLHGFHVHFGDNTAGCTS 60
QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDHCHIIIGRTLTV 120
QY 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154
DB 121 HEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ 154

RESULT 4

US-08-440-519-12

Sequence 12, Application US/08440519
Patent No. 5712087
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,519
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-519-12

Query Match 100.0%; Score 821; DB 1; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||

QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
|||||
DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
|||||

RESULT 5
US-08-440-519-12
Sequence 12, Application US/08440549
Patent No. 6312889
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation

STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,549
FILING DATE: 12-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-549-12

Query Match 100.0%; Score 821; DB 4; Length 1021;
Best Local Similarity 100.0%; Pred. No. 9.9e-84;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||
DB 1 MATKAVCVLKGDGPVQGIINFEQKESNGPVKVMGSIKGLTEGLHGFHVHFGDNTAGCTS 60
|||||

QY 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||
DB 61 AGPHFNPLSRKHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLIV 120
|||||

QY 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
|||||
DB 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
|||||

RESULT 6
US-08-679-493A-204
Sequence 204, Application US/08679493A
Patent No. 6303295
GENERAL INFORMATION:
APPLICANT: Taylor, Ethan W.
TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
FILE REFERENCE: 55-95
CURRENT APPLICATION NUMBER: US/08/679,493A
CURRENT FILING DATE: 1996-07-12
PRIOR APPLICATION NUMBER: 60/001203
PRIOR FILING DATE: 1995-07-14
PRIOR APPLICATION NUMBER: 60/003,112
PRIOR FILING DATE: 1995-09-01
NUMBER OF SEQ ID NOS: 216
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 204
LENGTH: 153
TYPE: PRT
ORGANISM: Homo sapiens
US-08-679-493A-204

Query Match 99.4%; Score 816; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.8e-84;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2	ATKACVCLKGGDPVGGIINFEQESNGPKVKGISIKGLTEGLHGFHVHFECDNTAGCTSA	61
Db	1	ATKACVCLKGGDPVGGIINFEQESNGPKVKGISIKGLTEGLHGFHVHFECDNTAGCTSA	60
Qy	62	GPHPNPLSRKGGPKDBERHVGDLGNNVTADKGVADVSIEDSVISLSDGHCIIIGRTLNVH	121
Db	61	GPHPNPLSRKGGPKDBERHVGDLGNNVTADKGVADVSIEDSVISLSDGHCIIIGRTLNVH	120
Qy	122	EXADDLGKGGNEESTKTGNAGSLACGVTGICIAQ	154
Db	121	EXADDLGKGGNEESTKTGNAGSLACGVTGICIAQ	153

RESULT 7

```

RESOLUT /
5290690-6
; Patent No. 5290690
; APPLICANT: MRABET, NADIR; LASTERS, IGNACE; S
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX, WILHE
; TITLE OF INVENTION: METHODS AND MEANS FOR
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO:6:
; LENGTH: 153
5290690-6

```

Query Match	99.4%;	Score 816;	DB 6;	Length 153;
Best Local Similarity	100.0%;	Pred. No. 2.8e-84;		
Matches 153:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	2	ATKAVCVLKGDGPVQGIINFEQKESGPNVYKWSIKGLTEGLGHGFHVFHFGDNTAGCTSA	61
Db	1	ATKAVCVLKGDGPVQGIINFEQKESGPNVYKWSIKGLTEGLGHGFHVFHFGDNTAGCTSA	60
Qy	62	GFHFNPLSRKHGGPKDERHVGLGNWTADKQGVADVISEDVSLSDGHCHIGRTLWVH	121
Db	61	GFHFNPLSRKHGGPKDERHVGLGNWTADKQGVADVISEDVSLSDGHCHIGRTLWVH	120
Qy	122	EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ	154
Db	121	EKADDLGKGGNEESTKTGNAGSLACGVIGIAQ	153

RESULTS

```

; Patent No. 5290690
; APPLICANT : MRABET, NADIR; LASTERS, IGNACE; STANSSENS, PATRICK
; MATTHYSSENS, GASTON; WODAK, SHOSHANA; QUAX, WILHELMUS J.
; TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
; STABILITY OF PROTEINS
; NUMBER OF SEQUENCES: 22
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/398,706
; FILING DATE: 25-AUG-1989
; SEQ ID NO:5:
; LENGTH: 154

```

Query Match 98.9%; Score 812; DB 6; Length 154;
Best Local Similarity 99.4%; Pred. No. 7.9e-84;
Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1	MATKAVCVLKGDPVQGIINFEQESGNPVKVMGSIKGLTEGLGHFHVHEFGDNTAGCTS	60
Qy		
1	MATKAVCVLKGDPVQGIINFEQESGNPVKVMGSIKGLTEGLGHFHVHEFGDNTAGCTS	60
Db		
61	AGPHNPILSRKHGPKDEERHVGDLGNVTADKGVADVYSIEDSVISLGGHICIGRTLLV	120
Qy		
61	AGPHNPILSRKHGPKDEERHVGDLGNVTADKGVADVYSIEDSVISLGGHICIGRTLLV	120
Db		

Qy 121 HEKADDLGKGNEESTKTGNAGSRLACGVIAQ 154
 |||||
Dd 121 HEKADDLGKGNEESTKTGNAGSRLACGVIAQ 154

RESULT 9

```

US-08-722-050-7
: Sequence 7, Application US/08722050
: Patent No. 5871729
: GENERAL INFORMATION:
: APPLICANT: YU, GUO-LIANG
: APPLICANT: ROSEN, CRAIG A.
: APPLICANT: FRASER, CLAIRE M.
: APPLICANT: GOCAYNE, JEANNINE D.
: TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
: STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
: CITY: WASHINGTON
: STATE: DC
: COUNTRY: USA
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/722,050
: FILING DATE: 23-JAN-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/225,757
: FILING DATE: 11-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: STEFFE, ERIC K.
: REGISTRATION NUMBER: 36,688
: REFERENCE/DOCKET NUMBER: 1488.1020001/EKS/AJK
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 152 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-722-050-7

```

Query Match	98.3%	Score 807;	DB 2;	Length 152;
Best Local Similarity	99.3%;	Pred. No. 2.8e-83;		
Matches 151: Conservative	0;	Mismatches 1;	Indels	0;
				Gaps

Qy	3	TRAVCVLKGDPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFVHVEFGDNTAGCTSAG	62
Db	1	TRAVCVLKGDPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFVHVEFGDNTAGCTSAG	60
Qy	63	PHFNPLSRKHGPGDDEERHVGDLGNTADKGVADVSIEDSVISLGDHCCIIGRTLVVHE	122
Db	61	PHFNPLSRKHGPGDDEERHVGDLGNTADKGVADVSIEDSVISLGDHCCIIGRTLVVHE	120
Qy	123	KADDLKGCGNEESTKTGNAGSRLACGVIGIAQ	154
Db	121	KADDLKGCGNEESTKTGNAGSRLACGVIGIIQ	152

RESULT 10

RESOLUT
5196335-1
; Patent No. 5196335
; APPLICANT: GRONER, YORAM

Db 1 MATNPVCLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 PGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152

RESULT 13

US-08-709-177-86
; Sequence 86, Application US/08709177
; Patent No. 5885799
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI LIM
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,177
; FILING DATE: 06-SEP-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/680,296
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 22300-20100.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-709-177-86

Query Match 97.1%; Score 797; DB 2; Length 841;
Best Local Similarity 98.0%; Pred. No. 3.8e-81;
Matches 149; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATKAVCLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Db 1 MATNPVCLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 61 PGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152
Db 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGI 152

RESULT 14
5252476-5
; Patent No. 5252476
; APPLICANT: HALLEWELL, ROBERT A.; MULLENBACH, GUY T.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE CLONING AND
; EXPRESSION IN MICROORGANISMS
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/222,352
; FILING DATE: 20-JUL-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 931,920
; FILING DATE: 14-NOV-1986
; APPLICATION NUMBER: 609,412
; FILING DATE: 11-MAY-1984
; APPLICATION NUMBER: 538,607
; FILING DATE: 03-OCT-1983
; SEQ ID NO: 5
; LENGTH: 144
5252476-5

Query Match 94.0%; Score 772; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.3e-79;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSAGPHFNPLSR 70
Db 1 GDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTSAGPHFNPLSR 60
Qy 71 KHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVHEKADDLGKG 130
Db 61 KHGGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTVHEKADDLGKG 120
Qy 131 GNEESTKTGNAGSLACGVIGIAQ 154
Db 121 GNEESTKTGNAGSLACGVIGIAQ 144

RESULT 15
5171680-3
; Patent No. 5171680
; APPLICANT: MULLENBACH, GUY T.; HALLEWELL, ROBERT A.; VALEZUELA,
; PABLO
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE ANALOGS HAVING NOVEL
; BINDING PROPERTIES
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/561,442
; FILING DATE: 01-AUG-1990
; SEQ ID NO: 3
; LENGTH: 152
5171680-3

Query Match 92.3%; Score 758; DB 6; Length 152;
Best Local Similarity 94.8%; Pred. No. 9.3e-78;
Matches 146; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MATKAVCLKGDGPVQGIINFEQESNGPVKVGSIKGLTEGLHGFHVEFGDNTAGCTS 60
Db 1 MATKAVAVLKGDGPVQGIINFEQ- -GDTVKVWGSIKGLTEGLHGFHVEFGDNTAGCTS 58
Qy 61 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 120
Db 59 AGPHFNPLSRKHGPKDEERHVGDLGNVTADKGVADVSIEDSVISLSDGHCIIIGRTLTV 118
Qy 121 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 154
Db 119 HEKADDLGKGGNEESTKTGNAGSLACGVIGIAQ 152

Search completed: January 28, 2003, 14:21:42
Job time : 5.19694 secs

QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSIQHSRRRTQFYFWLLNVLLGL 297
|||||
Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSIQHSRRRTQFYFWLLNVLLGL 300
QY 298 LVPVEDEHSTLLILGLVLLTLRLVPLVLLQOVKDTSLKSGFVTRKEMEVSFAEOLVQVY 357
Db 301 LVPVEDEHSTLLILGLVLLTLRLVPLVLLQOVKDTSLKSGFVTRKEMEVSFAEOLVQVY 360
QY 358 ELTLHHTQHDHNVVYTGAELELQQLFRYPPELLOTLTAVGGIGOLTAAKESGGRSRSG 417
Db 361 ELTLHHTQHDHNVVYTGAELELQQLFRYPPELLOTLTAVGGIGOLTAAKESGGRSRSG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESDSRSDVSSSALTATSVKDEISGEL 477
Db 421 SIVELIAGGSSCPVLSRKQKGVLLGEEEALEDDESDSRSDVSSSALTATSVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDITTEQPSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDITTEQPSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDPAMDNDGTQASSPISDSSQTTTEGPDASVTPSDSSSEIVLDGTDNOYVLGLQIGQP 597
Db 541 AVPSDPAMDNDGTQASSPISDSSQTTTEGPDASVTPSDSSSEIVLDGTDNOYVLGLQIGQP 600
QY 598 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 660
QY 658 QENKPCRIGDITGQSTDDSDAPLVHCVRLLSASFLTTGCKNVLPDRVRSVKALALSC 717
Db 661 QENKPCRIGDITGQSTDDSDAPLVHSVRLSASFLTTGCKNVLPDRVRSVKALALSC 720
QY 718 VGAVALHPESFEKLYKVPLODTEYPEQYVSDILNYIDHGDPQVRGATAILCGLTILCS 777
Db 721 VGAVALHPESFEKLYKVPLODTEYPEQYVSDILNYIDHGDPQVRGATAILCGLTILCS 780
QY 778 ILRSRFRHVGDMGTITRTLTGNTFTSLADCIPLLRKTLKDESSVTKLACTAVRNCVMSLC 837
Db 781 ILRSRFRHVGDMGTITRTLTGNTFTSLADCIPLLRKTLKDESSVTKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQILIDVLTIRNSSYWLVRTELLETLAEIDFRLVSFLAEAKENLHRAHYT 897
Db 841 SSSYSELGLQILIDVLTIRNSSYWLVRTELLETLAEIDFRLVSFLAEAKENLHRAHYT 900
QY 898 GLKLGQERVLNNVYIHLGDEDPVRHVAAAASLIRLVKLFYKCDQGOADPVVAVARQDS 957
Db 901 GLKLGQERVLNNVYIHLGDEDPVRHVAAAASLIRLVKLFYKCDQGOADPVVAVARQDS 960
QY 958 SVYLLKLMHETOPPSHFSYSTITRIYRYGNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
Db 961 SVYLLKLMHETOPPSHFSYSTITRIYRYGNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
QY 1018 RALTFCCEALCLLSTAFVPCVTSWGLWGHGCVPLPSASDESRSKSCVGMATMTLLTLLSSAW 1077
Db 1021 RALTFCCEALCLLSTAFVPCVTSWGLWGHGCVPLPSASDESRSKSCVGMATMTLLTLLSSAW 1080
QY 1078 FPLDLTSAHODALITLACNLLAASAPKSLRSSWASEEAPAAATQBEVMPALGDRAVLPVY 1137
Db 1081 FPLDLTSAHODALITLACNLLAASAPKSLRSSWASEEAPAAATQBEVMPALGDRAVLPVY 1140
QY 1138 EQFLSHLLKVINITCAHVLDVAPGPAIKAAPLSTNPPSLPIRRKGKEPEGEQASVPL 1197
Db 1141 EQFLSHLLKVINITCAHVLDVAPGPAIKAAPLSTNPPSLPIRRKGKEPEGEQASVPL 1200
QY 1198 SPKKGSEASAAQRSDTSGPTVTSKSSSLGSPHYLPSYLLKHDVLKATHYKVTLDLQNL 1257
Db 1201 SPKKGSEASAAQRSDTSGPTVTSKSSSLGSPHYLPSYLLKHDVLKATHYKVTLDLQNL 1260
QY 1258 STEKFGFLRSALDVLQSILELATQDICKVCVEEILGYLKSDFSREPMMATVCVQOOLLKT 1317
Db 1261 STEKFGFLRSALDVLQSILELATQDICKVCVEEILGYLKSDFSREPMMATVCVQOOLLKT 1320

QY 1318 LFGTNLASQFDGLSSNPSKSGQRAQRILGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
|||||
Db 1321 LFGTNLASQFDGLSSNPSKSGQRAQRILGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1380
QY 1378 VOAEQENDTSGWFDVLOKYSTQLKTNLTSTVTKNRADKNAIHNHRLFEPLVIRKALKQYTT 1437
Db 1381 VOAEQENDTSGWFDVLOKYSTQLKTNLTSTVTKNRADKNAIHNHRLFEPLVIRKALKQYTT 1440
QY 1438 TTCVOLQKQVLDLALQVLQRVNYCCLDSQVFTGFYKQFEYIEVQGFRESBAIIPNIF 1497
Db 1441 TTCVOLQKQVLDLALQVLQRVNYCCLDSQVFTGFYKQFEYIEVQGFRESBAIIPNIF 1500
QY 1498 FFLVLLSYERYHSKQIIGIPKIIQLQCDGIMASGRKA 1533
Db 1501 FFLVLLSYERYHSKQIIGIPKIIQLQCDGIMASGRKA 1536
RESULT 2
149729
HD protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I49729; I49730
R:Lin, B.; Nasir, J.; MacDonald, H.; Hutchinson, G.; Graham, R.; Rommens, J.M.; Hayde
Hum. Mol. Genet. 3, 85-92, 1994
A:Title: Sequence of the murine huntington disease gene: evidence for conservation an
A:Reference number: I49729; MUID:94214482; PMID:8162057
A:Accession: I49729
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3119 <RES>
A:Cross-references: GB:L23312; NID:G438804; PIDN:AAA37799.1; PID:G438805
A:Accession: I49730
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1521,2002-3119 <RE2>
A:Cross-references: GB:L23313; NID:G438806; PIDN:AAA37800.1; PID:G438807
C:Keywords: alternative splicing

Query Match 89.7% Score 7077.5; DB 2; Length 3119;
Local Similarity 90.4%; Pred. No. 0;
Matches 1387; Conservative 47; Mismatches 79; Indels 21; Gaps 4;

QY 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPPPPPP 60
Db 1 MGTEKLMKAFESLKSFOQQQQQQQQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 48
QY 61 LPQPP 120
Db 49 -----PP 100
QY 121 QKLLGIAMFLFLLCSDDAEDSVRMVADECLNKVTKALMDSNLPRLQLELYKEIKKNGCAPR 180
Db 101 QKLLGIAMFLFLLCSDDAEDSVRMVADECLNKVTKALMDSNLPRLQLELYKEIKKNGCAPR 160
QY 181 SLRAALNRFAELAHVLPQKCRPYLVNLLPCLTSTRKRPESVOETLAAAVPKIMASFGN 240
Db 161 SLRAALNRFAELAHVLPQKCRPYLVNLLPCLTSTRKRPESVOETLAAAVPKIMASFGN 220
QY 241 FANDNEIKVLLKAFIANLKSSSPTRTAAGSAVSIQHSRRRTQFYFWLLNVLLGLLVP 300
Db 221 FANDNEIKVLLKAFIANLKSSSPTRTAAGSAVSIQHSRRRTQFYFWLLNVLLGLLVP 280
QY 301 VEDEHSTLLILGLVLLTLRLVPLVLLQOVKDTSLKSGFVTRKEMEVSFAEOLVQVYELT 360
Db 281 MEDEHSTLLILGLVLLTLRLVPLVLLQOVKDTSLKSGFVTRKEMEVSFAEOLVQVYELT 340
QY 361 LHHTQHDHNVVYTGAELELQQLFRYPPELLOTLTAVGGIGOLTAAKESGGRSRSGIV 420
Db 341 LHHTQHDHNVVYTGAELELQQLFRYPPELLOTLTAVGGIGOLTAAKESGGRSRSGIV 400
QY 421 ELIAGGSSCPVLSRKQKGVLLGEEEALEDDESDSRSDVSSSALTATSVKDEISGELAAS 480
|||||

Db 477 SKAVATWHANTERQKQKETERIEKERMRLMAEDSEYKRLID-----QKDKRRLA 527
 QY 373 TGALELLOQLFRTPPPPELLQTL-AVGGIGQLTAKEEGSGRSRSISVELIAGGSSCS 431
 Db 528 -----YLLQO-----TDEYVANLJNLVWEHKQAQAAKRRRKKKBAENAEAGGSGALG 578
 QY 432 P-----VLSRKQKQKVLGEEBALEDDSESRSDVSSSALTASVKDEISG 475
 Db 579 PDGEIDESSQMSDLPVKVYTHTTGKVLFGPEAP-----KASQLDLAWLENNPGY 627
 QY 476 ELAASSGVTPGSAGHDITEQPRSOHTLOADSVDLASCDLTSSATGDEEDILSHSSSQ 535
 Db 628 EVA-----PRS-----DSESDSDYEEDEEERSSRQ 654
 QY 536 -----VSAPSPDAMDNDGTO-----ASSPISD--SSQTTTEGPDASVTPSDS----- 577
 Db 655 ETEKILLDPNSESEVSEKAKIETAKQDVDEYSQMSARGSQSYTTVAHAISERVEK 714
 QY 578 -SEIVLDGTDNOY--LGLQTGQPQDEDEEATGILPDE 611
 Db 715 QSALLINGTLKHQLOGLE-WWVSLYNNLNGILADE 750

RESULT 6
 A12043
 hypothetical protein alr1903 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp.
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: A12043
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih
 Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: A12043
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1547 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA073602.1; PID:g17130993; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr1903

Query Match 2.8%; Score 220.5; DB 2; Length 1547;
 Best Local Similarity 19.5%; Pred. No. 0.00041;
 Matches 237; Conservative 185; Mismatches 412; Indels 379; Gaps 54;

QY 81 EPLHPRKELSATKKDRVNHCLTICENIVAQSVRNSPEFQKLLGTAMELFLLCSDDAES 140
 Db 355 DEGFNPRDQASBSLKDLENHDL-----IQLAANDQIEF 388
 QY 141 DVMWAD-----ECLNKKVIALMDSNLPR-----LQLELYEIKKNGAPRSLR 183
 Db 389 RHQIQEYVTAECLLKLLPSDDSLQREYLNLYLKWTEPKLMLLELVD-----R 438
 QY 184 AALWRPAELAHVLPQKCPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGNFAN 243
 Db 439 ALAVRVNALKVDYO-----LGAKLAGAAKPE--LQOQTVDLIPGL----- 478
 QY 244 DNEIKVLLKAFITANKSSPTIRRTAAGSAVSICQHSRRRTQYFYSWLLNVLGLLVPVED 303
 Db 479 --KIPLLFKICLLGLTKSEKAISH-----LNPLL-----ED 507
 QY 304 EUSTL-----LILGLV---LTRYLVPLLOQVQKDTSLKSGVGRKMEVSPSAEQVLQ 355
 Db 508 ENSSVNDASAALGEIKSEAAIPGLTKLLEH--NSSVRYSAASALGEIKSEAAIPGLIN 565
 QY 356 VYELTLHHHTQHODHNVVTGALELLOQL-FRTPPPELLQTL-----TAVGGIGOLTA 405
 Db 566 LLE-----HENFDVRYRAASALGEIKSEAAIPGLINLLEHENS SVRYRAASALGQIKS 618

QY 406 AKEESGGRSGSISVELIAGGSS-----CSPVLSRKQKQKVLGEEBALEDDSESRSDVS 461
 Db 619 EQAIIIG-----LKLLEHENS SVRYRAASALGQIKSEAAIPGLIKLLEHEN-----SDVR 668
 QY 462 SSALTASVKEISGELAAASSGV-----TPGSAGHDITTEQ--PRSOHTLOADS 508
 Db 669 YRA--ASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINLLEHEN 726
 QY 509 VDL-----ASCDLTSSAT-----DGEEDILSHSSSOVSAPSPDAMD--LNDGTQ 552
 Db 727 FDVRYSAASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINLLEH 786
 QY 553 ASSPISDSQTTT-EGPDSAVTPSSSEIVLDGTDNQYLQ-ICQPODED--EATGIL 608
 Db 787 ENFDVRYSAASALGEIKSEAAIPGLINLLEHENFDVRYSAASALGEIKSEAAIPGLINL 846
 QY 609 PDEASEAFNFSMALQO-----AHLKKN-MSHCRQPSDSDVKFVLRLDATEPG- 656
 Db 847 EDENS DVRYSAASALGEIKSEAAIPGLINLLEHENS PVRRAASALGO--IKSEAAIPGL 904
 QY 657 ----DOENKPCR-----TKGDIGOSTDDSDAPLVHCVRLLSASF----- 692
 Db 905 INLLEHENS PVRRAASALGQIKSEAAIPGLINLLEHENS-----VRYSAASSLGQIKS 959
 QY 693 ---LTGGKNVLVPORFVRSVKALASCVAALHPESPFSKLYKVPDLTTEPPEQYV 749
 Db 960 EQAIIIGLKLK---EDENSSVRRRAASALG-----ETKSEQAI 994
 QY 750 SDILNYIDHGPPQVRGATAILCGTILCSILSRSFHVGDMMGTITLTGNTFSLADCIPL 809
 Db 995 PGLIKLLEHEDSDV-----RYRAASALGEIKS-----EQAI PG 1027
 QY 810 LRKTLKDESSVTCKLACTAVRNCVNSLCSSSYSELGLQIILVTLRNSVYMLVTELLE 869
 Db 1028 LTKLEDEDSD-----VRYRAASALGEIKSEAAIPGLINLLEHEDSS---VRYRAAS 1076
 QY 870 TLAEDIFR-----LVSELEKAEANLHRCAAHYTGLLKQERVLNNVTHLLGDEDPVRH 924
 Db 1077 TLGQIKSEAAIPGLINLLEHEDSKSVRDSAAASALGEIKSEAAIPG--LINLLEHENS SVR 1134
 QY 925 VAAASLIRLVKPKFYCKDQGOADPVAVARDQSSVYKLLMHETQPPSHFSVSTITRIY 984
 Db 1135 SAASALGEI-----KSEQAIPMLNRLNEEFVSAANNNTLYS 1172
 QY 985 GYNLLPSITDVTMEN-NLSRVIAVASHLITSTTRALTFCCEALCLLSTAFPCVWSLG 1043
 Db 1173 ALEALEAI-----QENCQYIRCLTEQKQLYLPSPQ-----SKTSLMVLHLS 1217
 QY 1044 WHCGVPPLSASDESCKCTVGMATMILTLLSSAW---FPLDLS-----AHODALILAGNL 1095
 Db 1218 LHFCTP-----DEANK-----WSNQLAMDQKQELQIPHLDALILSGDI 1255
 QY 1096 LAASAP-----KSLRSSWASEEENPAATKQEEV-WPALGDRALVPMVQLFSLH 1144
 Db 1256 ANKSTPDEYAAQAQFLDELRODFSLKFEQIIIAPNHDLNQNQISEEGYIPTLRKKQES 1315
 QY 1145 LKVINICAHVLDD 1157
 Db 1316 MD--ESC--VIDD 1324

RESULT 7
149681

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 03-Jun-2002
 C:Accession: I49681; I49046
 R:Welch, J.E.; Schatte, B.C.; O'Brien, D.A.; Eddy, E.M.
 Biol. Reprod. 46, 869-878, 1992
 A:Title: Expression of a glyceraldehyde 3-phosphate dehydrogenase gene specific to mo
 A:Reference number: I49681; MUID:92273722; PMID:1375514
 A:Accession: I49681
 A>Status: preliminary; translated from GB/EMBL/DBDJB

Tue Jan 28 14:23:37 2003

us-09-904-987-7.rpr

Page 10

Search completed: January 28, 2003, 14:20:44
Job time : 71.4611 secs

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %			ID	Description
		Match	Length	DB		
1	7821.5	99.1	3144	1	HD_HUMAN	P42858 homo sapien
2	7136.5	90.4	3119	1	HD_MOUSE	P42859 mus musculus
3	7045.5	89.3	3110	1	HD_RAT	P51111 rattus norv
4	5315	67.3	3148	1	HD_FUGRU	P51112 fugu rubrip
5	235	3.0	875	1	Y066_NPYOP	K89349 orgyia pseu
6	215.5	2.7	440	1	G3PT_MOUSE	Q64467 mus musculus
7	214	2.7	450	1	SN22_HUMAN	P51531 homo sapien
8	206	2.6	1822	1	2AP3_HUMAN	P49750 homo sapien
9	203	2.6	1206	1	FM14_MOUSE	Q05859 mus musculus
10	199.5	2.5	141	1	YPRO_OWFEU	P21260 owenia fusi
11	197	2.5	901	1	Y298_HUMAN	O15016 homo sapien
12	195.5	2.5	3703	1	ABF1_HUMAN	Q15911 homo sapien
13	194	2.5	1468	1	FNW1_MOUSE	Q05860 mus musculus
14	193.5	2.5	1365	1	SUZ2_DROME	P25172 drosophila
15	193.5	2.5	3726	1	ABF1_MOUSE	Q61329 mus musculus
16	189	2.4	1213	1	FNW1_CHICK	Q03957 gallus gall
17	188	2.4	485	1	SSGP_VOLCA	P21997 volvox cart
18	186	2.4	816	1	HUNB_DROVI	P13361 drosophila
19	185.5	2.4	1790	1	SEPA_EMENI	P78621 emericella
20	185.5	2.4	3149	1	TEGU_EBV	P03186 epstein-bar
21	183.5	2.3	4385	1	VF73_CAEL	Q09222 caenorhabdi
22	179.5	2.3	1192	1	RTN4_HUMAN	Q94qc3 homo sapien
23	179	2.3	1362	1	BED4_HUMAN	O60885 homo sapien
24	177.5	2.2	1443	1	E75C_DROME	P13055 drosophila
25	177	2.2	5430	1	ACF7_HUMAN	Q9upn3 homo sapien
26	176.5	2.2	487	1	EBN2_EBV	P12978 epstein-bar
27	176.5	2.2	1306	1	MSB2_YEAST	P32334 saccharomyc
28	176	2.2	2004	1	MOZ_HUMAN	Q92794 homo sapien
29	175.5	2.2	1849	1	T172_HUMAN	O14981 homo sapien
30	174	2.2	3164	1	TEGU_HSV1	P10220 herpes simp
31	173	2.2	1091	1	DIA_DROME	P48608 drosophila
32	171.5	2.2	2670	1	YAQ5_SCHPO	Q10105 schizosacch
33	171	2.2	905	1	SNF5_YEAST	P18480 saccharomyc

RP SEQUENCE OF 119-934 FROM N.A.
 RA Llovd C.; (APR-1995) to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 1212-1290 FROM N.A.
 RA Mungall A., Odell C.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1291-1860 FROM N.A.
 RA Mungall A.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 1862-2820 FROM N.A.
 RA Buck D.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE OF 2563-3144 FROM N.A.
 RC TISSUE=Frontal cortex, Brain, Retina, Caudate, and Muscle;
 RX MEDLINE=94093536; PubMed=7903579;
 RA Lin B., Rommens J.M., Graham R.K., Kalchman M., Macdonald H.,
 RA Nasir J., Delaney A., Goldberg Y.P., Hayden M.R.;
 RT "differential 3' polyadenylation of the Huntington disease gene
 results in two mRNA species with variable tissue expression.";
 RL Hum. Mol. Genet. 2:1541-1545(1993).
 RN [10]
 RP SUBCELLULAR LOCATION.
 RX MEDLINE=95375771; PubMed=7647777;
 RA Trottier Y., Devys D., Inbert G., Saudou F., An I., Lutz Y., Weber C.,
 RA Agid Y., Hirsch E.C., Mandel J.-L.;
 RT "Cellular localization of the Huntington's disease protein and
 discrimination of the normal and mutated form.";
 RL Nat. Genet. 10:104-110(1995).
 RN [11]
 RP CLEAVAGE BY APOPAIN.
 RX MEDLINE=96331285; PubMed=8696339;
 RA Goldberg Y.P., Nicholson D.W., Rasper D.M., Kalchman M.A., Koide H.B.,
 RA Graham R.K., Bromm M., Kazemi-Esfarjani P., Thornberry N.A.,
 RA Vaillancourt J.P., Hayden M.R.;
 RT "Cleavage of huntingtin by apopain, a proapoptotic cysteine protease,
 is modulated by the polyglutamine tract.";
 RL Nat. Genet. 13:442-449(1996).
 CC -1- FUNCTION: MAY PLAY A ROLE IN MICROTUBULE-MEDIATED TRANSPORT OR
 CC VESICLE FUNCTION.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN A VARIETY OF TISSUES WITH
 CC THE HIGHEST LEVEL EXPRESSION IN THE BRAIN (NERVE FIBERS,
 CC VARICOSITIES, AND NERVE ENDINGS). IN THE BRAIN THE REGIONS WHERE
 CC IT CAN BE MAINLY FOUND ARE THE CEREBELLAR CORTEX, THE NEOCORTEX,
 CC THE STRIATUM, AND THE HIPPOCAMPAL FORMATION.
 CC -1- PTM: CLEAVED BY APOPAIN DOWNSTREAM OF THE POLYGLUTAMINE STRETCH.
 CC THE RESULTING AMINO-TERMINAL FRAGMENT IS CYTOTOXIC AND PROVOKES
 CC APOPTOSIS.
 CC -1- POLYMORPHISM: THE POLY-GLN REGION OF HD IS HIGHLY POLYMORPHIC (10
 CC TO 35 REPEATS) IN THE NORMAL POPULATION AND IS EXPANDED TO ABOUT
 CC 36-120 REPEATS IN HD PATIENTS. THE REPEAT LENGTH USUALLY INCREASES
 CC IN SUCCESSIVE GENERATIONS, BUT CONTRACTS ALSO ON OCCASION. THE
 CC LONGER EXPANSIONS RESULT IN EARLIER ONSET AND MORE SEVERE CLINICAL
 CC MANIFESTATIONS OF THE DISEASE. THE ADJACENT POLY-PRO REGION IS
 CC ALSO POLYMORPHIC AND VARIES BETWEEN 7-12 RESIDUES. POLYGLUTAMINE
 CC EXPANSION LEADS TO ELEVATED SUSCEPTIBILITY TO APOPAIN CLEAVAGE AND
 CC LIKELY RESULT IN ACCELERATED NEURONAL APOPTOSIS.
 CC -1- DISEASE: DEFECTS IN HD ARE THE CAUSE OF HUNTINGTON'S DISEASE, AN
 CC AUTOSOMAL DOMINANT NEURODEGENERATIVE DISORDER CHARACTERIZED BY
 CC INVOLUNTARY MOVEMENTS (CHOREA), GENERAL MOTOR IMPAIRMENT,
 CC PSYCHIATRIC DISORDERS AND DEMENTIA. ONSET OF THE DISEASE OCCURS
 CC USUALLY IN THE THIRD OR FOURTH DECADE OF LIFE AND SYMPTOMS
 CC PROGRESSIVELY WORSEN LEADING TO DEATH IN 10 TO 20 YEARS. IT
 CC AFFECTS 1 IN 10,000 INDIVIDUALS OF EUROPEAN ORIGIN. NEUROPATHOLOGY
 CC OF HUNTINGTON'S DISEASE DISPLAYS A DISTINCTIVE PATTERN WITH LOSS
 CC OF NEURONS, SPECIALLY IN THE CAUDATE AND PUTAMEN (STRIATUM).
 CC -1- SIMILARITY: CONTAINS 10 HEAT REPEATS.
 CC -1- SIMILARITY: BELONGS TO THE HUNTINGTIN FAMILY.
 CC -1- DATABASE: NAME=HotMolBase; NOTE=HD entry;

CC WWW="http://bioinformatics.weizmann.ac.il/hotmolebase/entries/hunt1.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L12392; AAB38240.1; -;
 CC EMBL; L34020; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27350; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27351; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27352; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27353; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27354; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27355; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27356; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27357; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27358; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27359; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27360; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27361; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27362; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27363; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27364; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27365; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27366; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27367; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27368; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27369; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27370; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27371; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27372; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27373; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27374; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27375; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27376; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27377; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27378; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27379; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27380; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27381; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27382; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27383; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27384; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27385; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27386; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27387; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27388; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27389; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27390; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27391; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27392; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27393; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27394; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27395; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27396; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27397; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27398; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27399; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27400; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27401; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27402; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27403; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27404; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27405; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27406; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27407; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27408; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27409; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27410; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27411; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27412; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27413; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27414; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27415; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27416; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27417; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27418; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27419; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27420; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27421; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27422; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27423; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27424; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27425; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27426; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27427; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27428; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27429; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27430; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27431; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27432; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27433; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27434; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27435; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27436; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27437; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27438; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27439; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27440; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27441; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27442; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27443; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27444; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27445; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27446; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27447; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27448; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27449; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27450; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27451; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27452; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27453; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27454; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27455; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27456; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27457; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27458; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27459; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27460; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27461; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27462; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27463; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27464; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27465; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27466; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27467; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27468; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27469; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27470; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27471; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27472; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27473; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27474; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27475; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27476; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27477; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27478; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27479; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27480; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27481; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27482; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27483; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27484; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27485; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27486; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27487; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27488; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27489; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27490; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27491; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27492; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27493; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27494; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27495; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27496; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27497; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27498; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27499; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27500; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27501; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27502; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27503; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27504; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27505; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27506; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27507; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27508; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27509; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27510; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27511; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27512; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27513; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27514; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27515; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27516; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27517; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27518; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27519; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27520; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27521; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27522; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27523; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27524; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27525; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27526; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27527; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27528; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27529; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27530; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27531; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27532; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27533; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27534; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27535; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27536; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27537; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27538; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27539; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27540; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27541; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27542; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27543; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27544; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27545; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27546; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27547; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27548; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27549; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27550; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27551; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27552; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27553; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27554; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27555; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27556; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27557; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27558; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27559; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27560; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27561; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27562; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27563; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27564; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27565; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27566; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27567; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27568; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27569; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27570; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27571; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27572; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27573; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27574; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27575; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27576; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27577; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27578; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27579; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27580; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27581; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27582; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27583; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27584; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27585; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27586; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27587; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27588; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27589; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27590; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27591; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27592; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27593; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27594; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27595; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27596; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27597; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27598; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27599; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27600; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27601; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27602; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27603; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27604; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27605; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27606; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27607; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27608; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27609; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27610; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27611; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27612; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27613; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27614; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27615; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27616; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27617; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27618; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27619; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27620; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27621; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27622; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27623; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27624; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27625; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27626; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27627; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27628; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27629; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27630; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27631; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27632; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27633; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27634; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27635; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27636; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27637; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27638; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27639; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27640; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27641; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27642; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27643; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27644; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27645; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27646; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27647; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27648; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27649; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27650; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27651; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27652; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27653; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27654; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27655; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27656; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27657; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27658; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27659; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27660; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27661; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27662; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27663; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27664; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27665; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27666; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27667; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27668; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27669; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27670; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27671; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27672; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27673; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27674; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27675; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27676; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27677; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27678; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27679; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27680; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27681; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27682; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27683; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27684; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27685; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27686; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27687; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27688; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27689; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27690; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27691; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27692; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27693; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27694; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27695; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27696; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27697; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27698; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27699; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27700; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27701; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27702; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27703; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27704; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27705; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27706; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27707; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27708; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27709; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27710; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27711; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27712; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27713; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27714; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27715; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27716; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27717; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27718; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27719; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27720; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27721; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27722; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27723; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27724; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27725; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27726; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27727; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27728; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27729; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27730; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27731; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27732; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27733; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27734; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27735; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27736; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27737; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27738; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27739; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27740; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27741; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27742; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27743; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27744; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27745; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27746; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27747; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27748; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27749; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27750; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27751; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27752; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27753; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27754; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27755; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27756; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27757; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27758; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27759; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27760; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27761; -; NOT_ANNOTATED_CDS.
 CC EMBL; L27762; -; NOT_ANNOTATED_CDS.
 CC E

QY 298 LVPVEDEHSTLLILGLVLLTLRLVLPVLLQOQVKTSLKSGFVTRKMEYSPSAEQLVQVY 357
 Db 301 LVPVEDEHSTLLILGLVLLTLRLVLPVLLQOQVKTSLKSGFVTRKMEYSPSAEQLVQVY 360
 QY 358 ELTLHTHQHNDHNVVGALELQLQFLRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 417
 Db 361 ELTLHTHQHNDHNVVGALELQLQFLRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 420
 QY 418 SIVELLAGGSGSPVLSRKQKGVLLGEEALEDDSESRSVSSALTSYKDEISGEL 477
 Db 421 SIVELLAGGSGSPVLSRKQKGVLLGEEALEDDSESRSVSSALTSYKDEISGEL 480
 QY 478 AASSGVSTPCSGAGHDIITEQPSQSHQLQADSVDLASCDLTSSATDGEDILSHSSQVS 537
 Db 481 AASSGVSTPCSGAGHDIITEQPSQSHQLQADSVDLASCDLTSSATDGEDILSHSSQVS 540
 QY 538 AVPSDAMPDLNDGTQASSPISDSQTTTEGPDPSAVTPSDSSIVLDGTDNQYGLQIGQP 597
 Db 541 AVPSDAMPDLNDGTQASSPISDSQTTTEGPDPSAVTPSDSSIVLDGTDNQYGLQIGQP 600
 QY 598 QDEDEATGILPDEASEAFRNSMALQQAHLKKNKSHCHROPSDSSVDKFLRDEATEPGD 657
 Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKKNKSHCHROPSDSSVDKFLRDEATEPGD 660
 QY 658 QENKPCRIGDITGQSTDDSDAPLVHCVRILLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
 Db 661 QENKPCRIGDITGQSTDDSDAPLVHCVRILLSASFLLTGKKNVLPDRDVRVSVKALALSC 720
 QY 718 VCAAVALLHPESFSLKYKVPDLTTEYPEQYSDILNTIDHGDPOVRGATAILCGTLICS 777
 Db 721 VCAAVALLHPESFSLKYKVPDLTTEYPEQYSDILNTIDHGDPOVRGATAILCGTLICS 780
 QY 778 ILSRSRHFVGDWNGTIRLTGNTFSLADICPLLRKTLKDESSVTKLACTAVRNCVMSLC 837
 Db 781 ILSRSRHFVGDWNGTIRLTGNTFSLADICPLLRKTLKDESSVTKLACTAVRNCVMSLC 840
 QY 838 SSSYSELGLQLIIDVLTRNSSYWLVRTELLETAEIDFRLVFLKAEKAEHLRGAAHYT 897
 Db 841 SSSYSELGLQLIIDVLTRNSSYWLVRTELLETAEIDFRLVFLKAEKAEHLRGAAHYT 900
 QY 898 GLKQLQERVLLNNVTHLLGDEDPVRHVAASLRLVLPKLFYKCDQGDQADPVAVARDQS 957
 Db 901 GLKQLQERVLLNNVTHLLGDEDPVRHVAASLRLVLPKLFYKCDQGDQADPVAVARDQS 960
 QY 958 SYVLKLLMHETQPPSHFSVSTIRYGYNLLPSTDTVTMNNLSRVTAAYSHELITSTT 1017
 Db 961 SYVLKLLMHETQPPSHFSVSTIRYGYNLLPSTDTVTMNNLSRVTAAYSHELITSTT 1020
 QY 1018 RALTFGCCEALCLLSTAPPVCIWSLGNHCGVPPISASDESRSKCTVGMATMILTLLSSAW 1077
 Db 1021 RALTFGCCEALCLLSTAPPVCIWSLGNHCGVPPISASDESRSKCTVGMATMILTLLSSAW 1080
 QY 1078 FPLDLSAHODALILAGNLLAASAPKSLRSSWASSEENPAAKQEEVWPALGDRALVPMV 1137
 Db 1081 FPLDLSAHODALILAGNLLAASAPKSLRSSWASSEENPAAKQEEVWPALGDRALVPMV 1140
 QY 1138 EQLFSLHLKLVINICAHVLDVAVGPAIKAAPLSLTNPSPISPIRRKKEKPEGQASVPL 1197
 Db 1141 EQLFSLHLKLVINICAHVLDVAVGPAIKAAPLSLTNPSPISPIRRKKEKPEGQASVPL 1200
 QY 1198 SPKKGSEASASRSQSDTSGPVTTTSSKSSLSGFSYHLPVSKLHDVLCATHANYKVTLDLQN 1257
 Db 1201 SPKKGSEASASRSQSDTSGPVTTTSSKSSLSGFSYHLPVSKLHDVLCATHANYKVTLDLQN 1260
 QY 1258 STEKFGFLRSALDVLQILELATLQDICKVEELGVLKSCFREPDMATVCVQQLLKT 1317
 Db 1261 STEKFGFLRSALDVLQILELATLQDICKVEELGVLKSCFREPDMATVCVQQLLKT 1320
 QY 1318 LFGTNLASQFDGLSNFNSKSGRAQLGSSSVRPLGYHYCFMAYPTHFTQALADASLRNM 1377
 Db 1321 LFGTNLASQFDGLSNFNSKSGRAQLGSSSVRPLGYHYCFMAYPTHFTQALADASLRNM 1380
 QY 1378 VQAEQENDTSGWFDVLOKVSTQLKTNLTSTVTKNRADKNAINHRIPLFEPLVIALKQYTT 1437

Db 1381 VQAEQENDTSGWFDVLOKVSTQLKTNLTSTVTKNRADKNAINHRIPLFEPLVIALKQYTT 1440
 QY 1438 TTCVLOKQVLDLALQVLRVNYCLLSDQVFGVLKQFEYEVGQFRESEAIINIF 1497
 Db 1441 TTCVLOKQVLDLALQVLRVNYCLLSDQVFGVLKQFEYEVGQFRESEAIINIF 1500
 QY 1498 FFLVLLSYERYHSHKOITIGIPKIIQLCDGIMASGRKA 1533
 Db 1501 FFLVLLSYERYHSHKOITIGIPKIIQLCDGIMASGRKA 1536
 RESULT 2
 HD_MOUSE HD_MOUSE STANDARD: PRT: 3119 AA.
 ID AC P42859;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Huntingtin (Huntington's disease protein homolog) (HD protein).
 GN HD OR HDH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RC STRAIN=C57BL/6; TISSUE=Brain, and Spleen;
 RX MEDLINE=94214482; PubMed=8162057;
 RA Lin B., Nasir J., Macdonald H., Hutchinson G., Graham R.K.,
 RA Rommens J.M., Hayden M.R.;
 RT "Sequence of the murine Huntington disease gene: evidence for
 RT conservation, alternate splicing and polymorphism in a triplet (CCG
 RT repeat.";
 RL Hum. Mol. Genet. 3:85-92(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94278649; PubMed=80093370;
 RA Barnes G.T., Duyao M.P., Ambrose C.M., McNeil S., Persichetti F.,
 RA Srinidhi J., Gusella J.F., Macdonald M.E.;
 RT "Mouse Huntington's disease gene homolog (Hdh).";
 RL Somat. Cell Mol. Genet. 20:87-97(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95375771; PubMed=7647777;
 RA Trotter Y., Devys D., Imbert G., Saudou F., An I., Lutz Y.,
 RA Weber C., Agid Y., Hirsch E.C., Mandel J.L.;
 RT "Cellular localization of the Huntington's disease protein and
 RT discrimination of the normal and mutated form.";
 RL Nat. Genet. 10:104-110(1995).
 RN [4]
 RP SEQUENCE OF 1-181 FROM N.A.
 RX MEDLINE=95278941; PubMed=7759106;
 RA Lin B., Nasir J., Kaichman M.A., McDonald H., Zeisler J.,
 RA Goldberg Y.P., Hayden M.R.;
 RT "Structural analysis of the 5' region of mouse and human Huntington
 RT disease genes reveals conservation of putative promoter region and
 RT di- and trinucleotide polymorphisms.";
 RL Genomics 25:707-715(1995).
 CC -!- FUNCTION: MAY PLAY A ROLE IN MICROTUBULE-MEDIATED TRANSPORT OR
 CC VESICLE FUNCTION.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC short form; seem to be produced by alternative splicing. The
 CC short form cannot be explained by a simple splicing event.
 CC -!- TISSUE SPECIFICITY: THE HIGHEST LEVEL IS SEEN THROUGHOUT THE
 CC BRAIN, BUT IT IS ALSO FOUND IN THE STOMACH, HEART, TESTIS, ADIPOSE
 CC TISSUE, MUSCLE, SPLEEN, LIVER, AND KIDNEY.
 CC -!- DEVELOPMENTAL STAGE: PREDOMINANT EXPRESSION IN NEURONAL TISSUES AT
 CC ALL DEVELOPMENTAL STAGES. IN 14.5 DAY OLD EMBRYOS, IT IS ALSO
 CC DETECTED IN NON-NEURONAL TISSUES. THIS EXPRESSION IS DOWN-
 CC REGULATED IN LATER STAGES OF DEVELOPMENT.
 CC -!- POLYMORPHISM: THE FIRST POLY-PRO REPEAT STRETCH DIFFERS IN LENGTH

Query Match 2.6%; Score 206; DB 1; Length 1822;
Best Local Similarity 20.1%; Pred. No. 0.0073;
Matches 159; Conservative 93; Mismatches 310; Indels 228; Gaps 29;

QY 16 SFQQQQQQQQQQQQQQQQQQPPPPPPPPPPOLPPPQAQLLLPQPQPPPPPPPPPP 75
||||| ||||| :||| :||| ||||| ||||| ||||| |||||
DB 52 SFRHQHUAQLQGLQMHOKMQCVLQPHHLUPPPPLPPVPMGGGYGDMGPMPMPPPP 111
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 76 GPAAVEEPLHRPKKELSATKKDRNVNHCITICENTVAQSVRNSPEFQKLGTAMELFLLCS 135
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 112 GPALSYKQKQKYKHMLHHQRDGP-----PGLVPMELESPE-----S 149
QY 136 DDAESDVMMVADECLNKVIKALM-----DSNLPRLOLELYKEKKNGAPRSLRAALLWRF 189
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 150 PPVPDGSYMPSPQ-----SYMPPQPPPSYYPTSSQYLPPAQPSPSPSQSYLA 202
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 190 AELAHLVRPOKCRPYL---VNLLPCLTRTKRPESVQETLAAAVPKIMASFGFNANDNE 246
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 203 PTPSYSSSSSSSYLSHSQSYLSSQASPSPSGHSHKSOLLAPPSPAPPG-----NK 257
QY 247 IKVLKAFIANLKSSPTIRTAAGSAVSICQHRRRTQFYFSWLLNVLLGLLVPEDEHS 306
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 258 TTVOQEPLESGAKNKSTEQQAAPDPSTMTPEOQQQYWYR----- 299
QY 307 TLLILGVLLTLRYLVPLLQQ-----VKDTSLKGSEGT---RKMEVSPSA 350
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 300 -----QHLLSLQORTKVHLPGHKGVNAKDTPEPVKEVTYPATSQVPESPSS 348
QY 351 EQ-----LVQVELTLHTHQHDINV----- 371
DB 349 EEPPLPPNEEVPPPLPPEPOSEDEDARLKLQAAAAHWQHQQHVRGFQYGIMQK 408
QY 372 ---VTGALELLQQLERTPPPELLQTLTAVG-----GI-GOLTAAKEESSGRSRSGSI 419
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 409 HTQLQQILLQOYQQIIQ-PPPHIQATTPPGIPPPGVGPIPOLTAAPVPASSSQSQOV 467
QY 420 VEL-----IAGGG---SSCSVLSRKQKGVLLGEAELEDDESSDVSSSALTA 467
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 468 PEKPRPALLPVPVSGSAPPTYHPPL-----QSAGPSQVNSKAPLSKALPY 516
QY 468 SVKDISEIGELAASSGVSTPGSAGHD-----IITEQPRSQHTLOADS-----VD 510
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 517 SSFSDQGLGESSAAPSPITAVKMDPVRSGLLPPDRPSY-LESPPRGPRFDGPRREF 575
QY 511 LAS-CD-----LTSATQDGEDIEDLSH-----SSSOVASVPDSPAMDL----- 547
DB 576 LGSRCGPRPKPREGNPDGPRRYEGHPAEGTCKWGMIPRGPASQFYITPSTLSLP 635
QY 548 -NDGTQASPTSDDSQTTTEGPD SAVTP-SUSSETVLBGTNOYLGLOGQPDDEDEAT 605
DB 636 ROSGPQWKGPAPFAQQHQHQKPQSOAEPLSGNKEPLADTSSNQKNFKM-----QSAAF 689
QY 606 GTLPD-----EASEAFRNSSMALQQAHLK-----NMCHCRQPSDSSVBKFEVLDEEA 652
DB 690 STAADVKDKVAASQENENLSDSQEPPKSEVSGPVPENWDQNQVSMETQIDK---AQAV 746
QY 653 TEFGQENKP 662
DB 747 TQPVPLANKP 756

RESULT 9
FM14_MOUSE ID FM14_MOUSE STANDARD; PRT: 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Formin 1 isoform IV (Limb deformity protein).
GN FMN OR LD.
OS Mus musculus (Mouse).


```

165 QY LQLELYKEIK-----NGAPRSRAALWRFAEHLVROPKCRPYLVNLLPCLTRTSK-- 217
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
853 DB LHLEW-KDIOQAIFTVDDSVVDLETLAALYENRAQEDBLTKIRKY-----ETSKEE 903
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 QY -----RPESVQETLAAAVPKIMASGPNFANDNEIKVLKAF---IANLKSSPTIRRT 268
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
904 DB DLKLDKPEQFLHE-----LAQIPNFAERAQCIIFRAVSEGTSLSHRKVEIVTRA 954
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
269 QY AAGSAVSI COHSRRTQFYF5MLL-----NVLLG-----LLVPVEDEHS----- 306
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
955 DB SKG-----LLHMSVKDILALIFAGNYMGNGNRTGQADGYSLETLPLKLVKQYKSRDNGM 1007
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
307 QY TLLILGVLTLIRY-----LVPL-----LQOQVKDTSLKGSFGVTRKEMVSPS 349
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
1010 DB NLVDVYVKYLYRYDQEQAGTDKSVFPLPEPDQFFLASQVKFEDLLKDLKRLKQLEASQ 1067
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 QY AEQLV-----QVVELFLHH-----TOHQDHNVV7GALELLQLQFLRTPPPPELLQTLTAV 397
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
1070 DB QMKLVCKSPREVLQFPFKDLEEFFKAKKEHKEWESHLENAQKSE-----TTV 1117
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 QY GGTGQLTAAKEEGSGRRSGSIVELIAGGGSCS-----PVL5RKQKG-----KVLLGEEAL 450
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
1120 DB GYEGM-----KPKTGEKEVTPSYVFWV--WFECFSDPKTIWKRESKNISKERLKMQAQSVS 1177
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
451 QY EDDSESRSDVSSSALTASVKDEI 473
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :
1174 DB KLTSEKKVETKKINPTASLKERL 1196
   | | | | | : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10	
YPRO_OWEFU	
ID	YPRO_OWEFU
AC	STANDARD;
P21260:	P21261;
DT	01-MAY-1991 (Rel. 18, Created)
DT	01-MAY-1991 (Rel..18, Last sequence update)
DT	01-NOV-1995 (Rel. 32, Last annotation update)
DE	Hypothetical proline-rich protein (Fragment).

OS	Owenia fusiformis.
OC	Owenia
OC	Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OX	Sabellidae; Oweniidae; Owenia.
OX	NCBI_TaxID=6347;
[1]	
RN	SEQUENCE FROM N.A.
RP	MEDLINE=90147742; PubMed=2105723;
RX	Bakalara N., Collet J., Planells R., Thouveny Y., Fontes M.;
RT	'Presence in invertebrate genomes of sequences characterized by the
RT	repetition of the triplet CCGPurine.';
RL	Biochem. Biophys. Res. Commun. 166:66-73(1990).
KL	PIR; A34043; A34043.
KL	PIR; B34043; B34043.
KW	Hypothetical protein; DNA-binding.
DW	NON_TER 1
FT	DOMAIN 9 58 POLY-PRO.
FT	DNA_BIND 98 116 H-T-H MOTIF (POTENTIAL) .
FT	NON_TER 141 151
SQ	SEQUENCE 141 AA; 15745 MW; B294E884D152BDB9 CRC64;

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 14:04:57 ; Search time 77.9719 Seconds
(without alignments)
4077.505 Million cell updates/sec

Title: US-09-904-987-7
Perfect score: 7892
Sequence: 1 MATLEKLMKAFESLSKSPQQ.....DGIMASGRKASQPYRLCSP 1543

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	7811.5	99.0	3144	4 Q9UQB7	Q9UQB7 homo sapien
2	7041	89.2	3139	6 Q9CM99	Q9cm99 sus scrofa
3	5459.5	69.2	3121	13 Q42269	Q42269 brachydanio
4	674.5	8.5	3583	5 Q9V3N4	Q9v3n4 drosophila
5	668	8.5	3584	5 Q9U999	Q9u999 drosophila
6	358	4.5	75	6 O18897	O18897 canis famill
7	356.5	4.5	620	5 Q8R069	Q8r069 drosophila
8	294	3.7	1992	3 Q9P6T1	Q9p6t1 neurospora
9	293	3.7	67	11 Q9Z1I8	Q9z1i8 rattus norv
10	229.5	2.9	729	12 Q91TW1	Q91tw1 tupala herp
11	223	2.8	1214	4 Q8BQ04	Q8bq04 homo sapien
12	220.5	2.8	1547	16 Q8YV51	Q8yvs1 anabaena sp
13	210	2.7	192	4 Q90H11	Q90h11 homo sapien
14	210	2.7	5038	11 Q9QYX7	Q9qyx7 mus musculu
15	208	2.6	197	4 Q9UKC2	Q9ukc2 homo sapien
16	205	2.6	3536	5 Q9VZ30	Q9vz30 drosophila

17	205	2.6	5147	4 Q9Y6V0	Q9y6v0 homo sapien
18	203.5	2.6	506	5 Q961E9	Q961e9 drosophila
19	203.5	2.6	1185	11 Q8VDF4	Q8vdf4 mus musculu
20	203.5	2.6	1266	5 Q8SX65	Q8sx65 drosophila
21	202.5	2.6	243	5 Q8T1Y6	Q8t1y6 dictyostell
22	202	2.6	552	5 Q9GPB0	Q9gpb0 trypanosoma
23	201.5	2.6	1954	5 Q9VN82	Q9vn82 drosophila
24	201.5	2.6	2051	5 Q95TTS	Q95tt5 drosophila
25	200.5	2.5	780	11 Q9EQV7	Q9eqv7 rattus norv
26	200.5	2.5	858	11 Q99N37	Q99n37 rattus norv
27	199.5	2.5	1095	5 Q9BL72	Q9bl72 caenorhabdi
28	199	2.5	1893	5 Q24279	Q24279 drosophila
29	198	2.5	881	12 Q91GJ2	Q91gj2 epiphys po
30	198	2.5	1112	4 Q96PY5	Q96py5 homo sapien
31	197.5	2.5	1709	4 O15047	O15047 homo sapien
32	193	2.4	995	5 Q9V7E7	Q9v7e7 drosophila
33	192.5	2.4	752	11 Q8R506	Q8r506 rattus norv
34	192.5	2.4	830	11 Q99N38	Q99n38 rattus norv
35	192.5	2.4	4001	5 Q8WRQ7	Q8wrq7 drosophila
36	192	2.4	1795	4 Q9P2B8	Q9p2b8 homo sapien
37	191	2.4	599	10 P93797	P93797 volvox cart
38	190	2.4	1085	3 Q8TGW9	Q8tgw9 candida alb
39	189	2.4	409	10 Q9SBM1	Q9sbm1 volvox cart
40	189	2.4	432	11 Q9ESV6	Q9esv6 rattus norv
41	188.5	2.4	1209	11 Q924W6	Q924w6 mus musculu
42	188.5	2.4	1236	3 Q9C105	Q9c105 schizosacch
43	188	2.4	1368	5 Q9V6J0	Q9v6j0 drosophila
44	187.5	2.4	2025	11 Q99PP2	Q99pp2 mus musculu
45	187	2.4	773	10 Q8SY7Y	Q8sy7y oryza sativ

ALIGNMENTS

RESULT 1
Q9UQB7
ID Q9UQB7 PRELIMINARY; PRT; 3144 AA.
AC Q9UQB7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Huntingtin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20469406; PubMed=11013077;
RA Matsuyama N., Hadano S., Onoe K., Osuga H., Shouguchi-Miyata J.,
RA Gondo Y., Ikeda J.-E.;
RT "Identification and characterization of the miniature pig Huntington's
RT disease gene homolog: evidence for conservation and polymorphism in
RT the CAG triplet repeat.";
RL Genomics 69:72-85(2000).
DR EMBL: AB016794; BAA36753.1; .
DR InterPro: IPR000091; Huntingtin.
DR InterPro: IPR002965; P-rich_extensn.
DR Pfam: PF03541; Huntingtin; 1.
DR PRINTS: PR00375; HUNTINGTIN.
DR PRINTS: PR01217; PRICHEXTENS.
SQ SEQUENCE 3144 AA; 347839 MW; 3F2BFFFEFE8E5D8E CRC64;

Query Match 99.0%; Score 7811.5; DB 4; Length 3144;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 MATLEKLMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPP 60
QY 58 QPILLPQPQPP 117

D	b	1141	EQLFSHLLKVINICAHVLDVDVAPGPAITKAALPSLTNPSPISPIRRKCKEKEPGEQASVPL	1200
Q	y	1198	SPKKGEASAARSQSQTSGPVTTSSKSSILGSFYHLPYSYKLHDVYLKATHANYKYVTLDLQN	1257
D	b	1201	SPKKGEASAARSQSQTSGPVTTSSKSSILGSFYHLPYSYKLHDVYLKATHANYKYVTLDLQN	1260
Q	y	1258	STKFGGFURSALDVLUSQILELATIQTIGKCVEETILGYLKSCFSREPMMATVCVQQLLKT	1317
D	b	1261	STKFGGFURSALDVLUSQILELATIQTIGKCVEETILGYLKSCFSREPMMATVCVQQLLKT	1320
Q	y	1318	LFTGNLASQFDGLSSNPKSQSQAORGLGSSVRPGLYHYCFMAYPTHFTQALADASLRNM	1377
D	b	1321	LFTGNLASQFDGLSSNPKSQSQAORGLGSSVRPGLYHYCFMAYPTHFTQALADASLRNM	1380
Q	y	1378	VQAEQENDTSGWFDVLQKYSTQLTNLTSVTKNRADKNAINHNRLFEPLVIKALKQYTT	1437
D	b	1381	VQAEQENDTSGWFDVLQKYSTQLTNLTSVTKNRADKNAINHNRLFEPLVIKALKQYTT	1440
Q	y	1438	TTCVOLQKOVLDLAOLVQLRVNYCCLDSQVFIFGVULKQFEYTEVGQFRESEAIIPNIF	1497
D	b	1441	TTCVOLQKOVLDLAOLVQLRVNYCCLDSQVFIFGVULKQFEYTEVGQFRESEAIIPNIF	1500
Q	y	1498	FFVLVLSYERHSKQIIIGIPKIIQLCDGIMASGRKA	1533
D	b	1501	FFVLVLSYERHSKQIIIGIPKIIQLCDGIMASGRKA	1536
RESULT 2				
Q	G9CM99	ID	O9CM99 PRELIMINARY; PRT; 3139 AA.	
A	C	AC	O9CM99;	
D	T	DT	01-MAR-2001 (T-EMBLrel_16, Created)	
D	T	DT	01-MAR-2001 (T-EMBLrel_16, Last sequence update)	
D	T	DT	01-JUN-2002 (T-EMBLrel_21, Last annotation update)	
D	E	DE	Huntingtin.	
O	S	OS	Sus scrofa [pig].	
O	C	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
O	C	OC	Mammalia; Eutheria; Cetartiodactyla; Sulina; Suidae; Sus.	
O	X	OX	NCBI_TaxID=9823;	
R	N	RN	[1]	
R	P	RP	SEQUENCE FROM N.A.	
R	C	RC	STRAIN-CSK GOETTINGEN; TISSUE=TESTIS;	
R	X	RX	MEDLINE=20469406; PubMed=11013077;	
R	A	RA	Matsuyama N., Hadano S., Onoe K., Osuga H., Shouguchi-Miyata J.,	
R	A	RA	Gondo Y., Ikeda J.-E.;	
R	T	RT	"Identification and characterization of the miniature pig Huntington's	
R	T	RT	disease gene homolog: evidence for conservation and polymorphism in	
R	T	RT	the CAG triplet repeat.";	
R	L	RL	Genomics 69:72-85(2000).	
D	R	DR	EMBL; AB016793; BAA36752.1; -.	
D	R	DR	InterPro; IPR002106; AATRNA_ligaseII.	
D	R	DR	InterPro; IPR001092; HLH_basic.	
D	R	DR	InterPro; IPR000091; Huntingtin.	
D	R	DR	Pfam; PF03541; Huntingtin; 1.	
D	R	DR	PRINTS; PR00375; HUNTINGTIN.	
D	R	DR	PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.	
D	R	DR	PROSITE; PS00339; HELIX_LOOP_HELIX; UNKNOWN_1.	
S	Q	SQ	SEQUENCE 3139 AA; 344796 MW; 051D0119A727270F8 CRC64;	
Query Match				
Best Local Similarity 88.9%; Score 7041; DB 6; Length 3139;				
Matches 1371; Conservative				
Indels 16; Gaps				
Q	y	1	MATLEKLMKAFESLSKFQQQQQQQQQQQQQQQQQQQQQQPPPPPQLPQPQPPQAQPL	60
D	b	1	MATLEKLMKAFESLSKFQQQQQQQQQQQQQQQQQQQQQQPPPPPQPQPQPPQ--TQPPQ-----55	
Q	y	61	LPOQPPPPPPPPPPGPAVAEEPLHRPKELSATKDRVNHCLTICENIVAQSVRNSPEF	120
D	b	56	--ppppppppppppGPAVAEEPLHRPKELSATKDRVNHCLTICENIVAQSLRNSEP	113
Q	y	121	OKLLGLTAMELFILCSDDAESDVRYMADECLNKIVIKALMSDLNLRLQLJELYKETKNGAPR	180

Db 114 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIKALMDSNLPRLQLELYKEIKNGAPR 173

QY 181 SLRAALWRFELAHVLPKQCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGN 240

Db 174 SLRAALWRFELAHVLPKQCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGN 233

QY 241 FANDNEIKVLLKAFIAKLSSTPIRRTAAGSAVSTICQHSRRTQFYFWLLNLLGLLVP 300

Db 234 FANDNEIKVLLKAFIAKLSSTPIRRTAAGSAVSTICQHSRRTQFYFWLLNLLGLLVP 293

QY 301 VEDEHSTLLTLGLVLLTLRYLVLPQQVYKDTSLKSGFGVTRKREMEVSPSABQLVQVYELT 360

Db 294 VEDEHSTLLTLGLVLLTLRYLVLPQQVYKDTSLKSGFGVTRKREMEVSPSABQLVQVYELT 353

QY 361 LHRTQHDHNVVTGALLELQQLFRTPPPPELLQTLTAAGIGOLTAAKPEESGGRSRGSIV 420

Db 354 LHRTQHDHNVVTGALLELQQLFRTPPPPELLQTLTAAGIGOLTAAKPEESGGRSRGSIV 413

QY 421 ELIAGGSSCSPVLSRKQKGVLLGEEBALEDDSDRSVSSALTSVKDEISGELAAS 480

Db 414 ELIAGGSSCSPVLSRKQKGVLLGEEBALEDDSDRSVSSALTSVKDEISGELAAS 473

QY 481 SGVSTPGSA-----GHDIITEQPRSOHTLOADSVDLASCDLTSSATDCDEEDILSHSS 533

Db 474 SGVSTPGSAADSVGHDIITEQPRSOHTLOADSVDLSGCDLSSAATDCDEEDILSHSS 533

QY 534 SOVSAPVSDPAMDNDGTQASPISSDSTQTTTEGPDASVATPSDSSEIVLDGTNOYLGLO 593

Db 534 SOVSAPVSDPAMDNDGTQASPISSDSTQTTTEGPDASVATPSDSSEIVLDGTNOYLGLO 593

QY 594 IQOPO--DEDEATGILPDEASEAFRNSMALQQAHLKNNSHCRQPSDSVVKFVLRDE 651

Db 594 LGQOPDADADEDAAGLLPHGGPDAPRSPPTALQSHVLSKLSGHRSQPSDSVVKFVLRDE 653

QY 652 ATEPGDQENKPCRIGKIDTQSTDDSDAPLVHCVRLLSALLTGGKNVLPDRDVRVSVK 711

Db 654 AAEAGEPENKPCRIGKIDTQSTDDSDAPLVHCVRLLSALLTGGKNVLPDRDVRVSVK 713

QY 712 ALALSCVGAVALHPESFSLKLYVPLDTTEYPEEQVSDILNYIDHGDQVQRGATAILC 771

Db 714 ALALSCVGAVALHPESFSLKLYKAPLDSVEYDQEVSDILSYIHGGDPQVRGATAILC 773

QY 772 GTLICSILSRHFVHGMGTIRTLTGNFTSLADCIPLLRKTLKDESSVTKLACTAVRN 831

Db 774 GTLVSSILSRHFVHGMGTIRTLTGNFTSLADCIPLLRKTLKDESSVTKLACTAVRN 833

QY 832 CVMSCSSSYSELGLQILIDVITLRNSSYWLVRTELLETLAEIDFRLVSPLEAKAENLHR 891

Db 834 CVAALCGSSYSQWGLQITLDTLLTRSSSYWLVRTELLETLAEIDFRLVSPLEAKAENLHR 893

QY 892 GAHHYTGLLKQERVLNNVYHLLGDEDPVRHVAASLRVLPKLFYKCDQOQADPVVA 951

Db 894 GAHHYTGLLKQERVLNNVYHLLGDEDPVRHVAASLRVLPKLFYKCDQOQADPVVA 953

QY 952 VARDQSSVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTWENLRSVIAVSHS 1011

Db 954 VARDQSSVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTWENLRSVIAVSHS 1013

QY 1012 LITSTTRALTFCCEALCLLSTAFFVCVWSLGHGCVPLPLSADSESRKCTVGMATMIT 1071

Db 1014 LITSTTRALTFCCEALCLLSTAFFVCVWSLGHGCVPLPLSADSESRKCTVGMATMIT 1073

QY 1072 LLSAFAWPLDLSAHODALILAGNLLAASAPKSLRSWASEEENANPAATQOEVEWPAALGR 1131

Db 1074 LLSAFAWPLDLSAHODALILAGNLLAASAPKSLRSWASEEENANPAATQOEVEWPAALGR 1133

QY 1132 ALVPMVQBLFSLHLKVLNICAHVLDVADGPAIKAAALPSLTPPSPISPTRRKGEKEPEGE 1191

Db 1134 SLVPMVQBLFSLHLKVLNICAHVLDVADGPAIKAAALPSLTPPSPISPTRRKGEKEPEGE 1193

QY 1192 QASVPLSPKKGSGSASASROSDTSGPVVANKSSLSGSCFCHLPVSYLKLHDVLKATHANYKV 1251

Db 1194 QASVPLSPKKGSGSASASROSDTSGPVVANKSSLSGSCFCHLPVSYLKLHDVLKATHANYKV 1253

QY 1252 TLDLQNSTEKGFLRSALDVLQSLLELATLQDICKVCVEILGYLKSCFSREPMATVCV 1311

Db 1254 TLDLQSGTEKFRGLLSALDVLQSLLELATLQDICKVCVEILGYLKSCFSREPMATVCV 1313

QY 1312 QQLLKTFLGNTLASQDFGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAYTHFTQALAD 1371

Db 1314 QQLLKTFLGNTLASQDFGLSASTPSKCGRAQRLGSSSVRPGLYHYCFMAYTHFTQALAD 1373

QY 1372 ASURNVQAEOENDTSGWFDVLQKYSTQLKNTLTSVTNNRADKNAIHNRHILFEPLVKA 1431

Db 1374 ASURNVQAEOENDTSGWFDVLQKYSTQLKNTLTSVTNNRADKNAIHNRHILFEPLVKA 1433

QY 1432 LKQYTTTTCVQLQKQVLDLLAQLVLRVNYVCLLSDQVFIQFVLKQFYIEVGOFRESEA 1491

Db 1434 LKQYTTTTCVQLQKQVLDLLAQLVLRVNYVCLLSDQVFIQFVLKQFYIEVGOFRESEA 1493

QY 1492 IIPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1494 IIPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1535

RESULT 3

O42269

ID O42269 PRELIMINARY: PRT: 3121 AA.

AC O42269;

DT 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE HUNTINGTIN.

GN HD OR HD.

OS Brachydanio rerio (zebrafish) (zebra dario).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Karlovich C.A., Ramirez L., John R., Stainier D., Myers R.M.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 9-138 FROM N.A.

RA Williams L.C.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF052602; AAC63983.1; -

DR EMBL; AF052603; AAC63984.1; -

DR EMBL; AF028731; AAC05309.1; -

DR ZFIN; ZDB-GENE-990415-131; hd.

DR InterPro; IPR000091; Huntingtin.

DR Pfam; PF03541; Huntingtin.

DR PRINTS; PR00375; HUNTINGTIN.

SQ SEQUENCE 3121 AA; 346666 MW; 3CC884E199D324AC CRC64;

Query Match 69.2%; Score 5459.5; DB 13; Length 3121;

Best Local Similarity 68.6%; Pred. No. 0;

Matches 1075; Conservative 189; Mismatches 210; Indels 93; Gaps 11;

QY 1 MATLEKLMKAFESLKSFTQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPPP 60

Db 1 MATMEKLMKAFESLKSFTQQQQ----- 21

QY 61 LPQPP 120

Db 22 -----GPLSAELVQKQKDLATTKDRVTHCTICENTVAQSLRSPF 66

QY 121 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIKALMDSNLPRLQLELYKEIKNGAPR 180

Db 67 QKLGIAEMLFLLCSDDAESDVRWVADECLNKVIKALMDSNLPRLQLELYKEIKNGAPR 126

QY 181 SLRAALWRFELAHVLPKQCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGN 240

Db 127 SLRAALWRFELAHVLPKQCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGN 186

QY 241 FANDNEIKVLLKAFIANLKSSPTIRRTAGSVAISICQHSRRTOFYXWLLNVLLGLVP 300
Db 187 FANDGETKMLKAFVANLKSSPTIRRTAASSAVSVCQHSRRTHFYTWLLNVLLGLVVP 246
QY 301 VEDEHSTLLTGLVLLRLRYLPLVLLQOVKDTSLKSGFVTRKEMESPVAEOLVQVVELT 360
Db 247 VDEHSHLLTGLVLLRLRYLPLVLLQOVKDTSLKSGFVTRKEMESPVAEOLVQVVELT 306
QY 361 LHHTQHDHNVVTGALELLOQLFRTPELLOTLITAVGGIGQLTAAKESRSGRSSTV 420
Db 307 LHHTQHDHNVVTASLELLOQLFRTPELLOTLITAVGGIGQLTAAKESRSGRSSTV 366
QY 421 ELIAGGSSCPVLSRKOKKVLGGBEALDDSESDVSSSALITASVKDEISGELAAS 480
Db 367 ELIA-GGSTCPLLRKQKGLLSGEEGLDPE-RAEVTTGSFTASVGGSDSSSEAPS 424
QY 481 SGVSTPGSAGHDITEOPR-SOHTLO-ADSVLLASCDL---TSSATDGGDEEDILSHSSQ 535
Db 425 SGVSSLGTS--DIITEOPRSQHALQGDSDVLSASEQVGPDPDPDEEDMLSRSSG 482
QY 536 VSAYPS---DPAMDND---GTQASSPISDSQTTTEGPDSPAVTSDSSSIVLDGTDNQY 589
Db 483 GAGLVSTSGDLVTDANQMSAGAVSSSPSESSQTTTEGPDSPAVTSDCAELVLDGSESQY 542
QY 590 LGLOIQQPQDEDEATGILPDEASEAFRNSMALQOAHLLKNKSHCHOPSDSSVDKFEVLR 649
Db 543 SGMQIGTLQDEEEGSAAPPDPKPEPFSQSALASKPHLLLEGKGNKOSDSSVDRIPIK 602
QY 650 DEATEPGDOENKCRKIGDQSDDDSDAPLVHCVRLLSASELITGGKNVLPDROVRVS 709
Db 603 EEVLEPAELDNKPSRIKGDIGHYTDKPEELMHCVRLLSASELITGGKNVLPDNEVRVS 662
QY 710 VKALASCVGAVALHPESFSLKYLKVPDLTTEYPEBOYVSDILNYIDHGDQPVGATAI 769
Db 663 VKALAVSCVGAVALHPESFSLKYLKVPDLTTEYPEBOYVSDILNYIDHGDQPVGATAI 722
QY 770 LCGTILCSILSRFHVGDMGTTRTGTNTFSLADICPLRLTKDSESVTKCLACTAV 829
Db 723 LCGALIOAIQKTRNTETWLSQISQVTSQVSSVTLENFVPLLOSLKDESSVTCKMACAV 782
QY 830 RNCVNLSCSSYSELGLQIIVDITLNRNSVWLVRTELLETLABIDPRLVSFLAKAENL 889
Db 783 RHCIMALCNGSLSELGLQIIVDITLNRNSVWLVRTELLETLABIDPRLVSFLAKAENL 842
QY 890 HRGAHYTGLKLOPRLNVNVIHLGDDEPRVHVAASLIRLVKLFYKCDQGOQADPV 949
Db 843 HKGERHYTGLKLOPRLNVNVIHLGDDEPRVHVAASLIRLVKLFYKCDQGOQADPV 902
QY 950 VAVARDQSVVLKLMHETQPPSHFSVSTIIRYGVNLPSPITDVTMNNLSRVIAVS 1009
Db 903 VAIARDQSVVLKLMHETQPPSHFSVSTIIRYGVNLPSPITDVTMNNLSRVIAVS 962
QY 1010 HELITSTRALTFCCEALCLLSTAFFVCIVISLGHGCVPLS----- 1052
Db 963 HALTSSTSRAMTFCCEALCLLSTAFFVCIVISLGHGCVPLS----- 1022
QY 1053 -----ASDESRSKCTVGMATMILTLSSAWFPPLDLSAQHQAALILAGNLLAASAPKSLRS 1106
Db 1023 SLUSQSGSNEEVRSRTVGVASVNLSSLISSAWFPPLDLSAQHQAALILAGNLLAASAPKCKMS 1082
QY 1107 SWASEEANPAATKOEVPWALGDRALVPMVQELFSLHLLKVINICAHVLDVAVGPAIKA 1166
Db 1083 PNAGEESSPASSKVEEPWALGDRALVPMVQELFSLHLLKVINICAHVLDVAVGPAIKA 1142
QY 1167 ALPSLTNPSPSLPIRKKKEKEPGEQASVPLSPKKGSEASASROSDTSGPVTTSKSSSL 1226
Db 1143 SLPSLTNPSPSLPIRKKKEKEKDMGAGTTPMSPKKGSTN-TGRAADSTATAVANKSTTL 1201
QY 1227 GSFYHPLSVLKHLDVHAKATHYKVTLDLQNSTEFGFGLRSALDVLSQLLELATLQD 1286
Db 1202 GSFYHPLSVLKHLDVHAKATHYKVTLDLQNSTEFGFGLRSALDVLSQLLELATLQD 1261
QY 1287 KCVEEILGYLKSCFREPMMATVVCQQLKTLFGTNLASQFDGLSSNFSKSGRAQLRGS 1346

Db 1262 KCVEEILGYLKSCFREPMMATVVCQQLKTLFGTNLASQFDGLSSNFSKSGRAQLRGS 1321
QY 1347 SSVRGLYHYCFMAYPTHFTQALADASLRNVQAEQENDTSGWDFVQKYSTQLKTNLTS 1406
Db 1322 SSVRGLYHYCFMAYPTHFTQALADASLRNVQAEQENDTSGWDFVQKYSTQLKTNLTS 1381
QY 1407 VTKNADKNATHNHLRLEPLVIRKALKQYTTTTCVQLQKQVLDLLAQLVQLRVNVCLLDS 1466
Db 1382 VTKNADKNATHNHLRLEPLVIRKALKQYTTTTCVQLQKQVLDLLAQLVQLRVNVCLLDS 1441
QY 1467 DOVFICFVLFKOFEXIEVQGFRESEAIIPNIFFLVLLSYERYHSKQIITGPKIITOLCDGI 1526
Db 1442 DOVFICFVLFKOFEXIEVQGFRESEAIIPNIFFLVLLSYERYHSKQIITGPKIITOLCDGI 1501
QY 1527 MASGRKA 1533
Db 1502 MASGRKA 1508

RESULT 4
Q9V3N4
ID Q9V3N4 PRELIMINARY: PRT: 3583 AA.
AC Q9V3N4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Huntingtin protein
GN HUNTINGTIN OR HSP40/HD OR CG9905
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Bessley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Glodok A., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Harris N.L., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Pittman K.A., Nixon K., Nusskern D.R., Pacleb J.M., M.G.,
RA Palazzolo M., Pittman K.A., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirsky R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,


```

Db 1421 PPLAELDALATNSAPSYDPTSGSTSTSSSGSAAAVSAASAYFEASYGIGIAEG 1480
QY 986 YNL-LPSITD---VTMNNLSVIAAASHLITSTRALTTCGCCALCILLSTAPPVCITWS 1041
Db 1481 HVPALASASORQIAQEKVLAKVLRINKLMTLNDKNVQGIITVYALRLLRHFNFDYQ 1540
QY 1042 LGWHGCVPLPSASDESRSKCTVGMATMLTLLSSAWF-----PLDLSAHQDALILAGNLLA 1097
Db 1541 QAWL-----EENFVEICISYAYNNATAADLGCCNDLJDVNGKMA 1581
QY 1098 ASAPKSLRSSWASEBEANPAATKQEVVPALCDRALVPMVEOLFSLHLKVINICAHVDD 1157
Db 1582 -----GAMUSSGEPNTA-----HLDFLLRHVSVMKMLNIYVHLVTN 1615
QY 1158 VAPGAIPAALPSLTNPPLSPSIRRRKKEKEPEQASVPLPKKGSSEASASROSDTSGP 1217
Db 1616 QRPPTA-----GSGSGSSSKQPKSELFAREQPAAT--- 1646
QY 1218 VTTSSKSSGSPYHPLSVLKLHVLKATHANYKVTLDLQNSTEKFGFLRSALDVLSQL 1277
Db 1647 -----LQALGYFAGDVVYMKLYNLRGANDSYKITIN-QEAGSLICLLKTLCHAVSLCL 1700
QY 1278 E---LATLQDICKVEEILGYLKSCFSREPMMATVGVQOLKTLFTGTNLASQFDGLSNP 1334
Db 1701 EGMASAPPEL-KLIEEILHLVLRILNYAPAECAVCLRLKYLFAQNYASQ---VRLQP 1756
QY 1335 SKSOGRAORLAGSSVRPGLYHYCFMAYP-----THFTQALADASLRNM 1377
Db 1757 SAG-----IGNGTIG-G-HIAAFMRPFAAKGRGHGASSLTLLPTINSKPAVAVGSGORGA 1809
QY 1378 -VOAEQENDTSGFVQVLOKVTQKLTNTSVTKNRADKNAIHNIRLFEPLVIRKALKQYT 1436
Db 1810 PTDARQPIDAG---PLQDMGMLFVHGLQPPPTPAGDCVRL---IKLFEPMVYICLILFM 1862
QY 1437 TTTCVOLQKQVLDLALQVLQVRYNVCILDSQVFGFVLFKQFVIEVQGFRESEAITPNI 1496
Db 1863 KSNAL-VQAPILRLSLQDLNVTYILDSKNVIFDQVLSNMDLIEGGIDRNAFIMVPPM 1921
QY 1497 FFFVLVLSYRHSKQIIGIKPIIQLCDGIMASG 1530
Db 1922 LRFLVOLTHK--SDRLQITIPKISITNLLANG 1953

RESULT 6
O18897
ID O18897 PRELIMINARY; PRT; 75 AA.
AC O18897;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-JAN-1998 (Tremblrel. 05, Last annotation update)
GN HD.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RA Shiba H., O'Brien D.P., Chen Y.-W., Johnson G.S.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF029754; A0884207.1;
FT NON_TER 1
FT NON_TER 75
SQ SEQUENCE 75 AA; 8133 MW; ADD779011B8CD367 CRC64;

Query Match 4.5%; Score 358; DB 6; Length 75;
Best Local Similarity 94.7%; Pred. No. 3,9e-15;
Matches 71; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 670 GQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSCVGAVALHPESF 729
Db 1 GQSTDDSDVPLVHCVRLLCASFLLTGKKNVLPDRDVRVSVKALALSCVGAVALHPESF 60

```

```

QY 730 FSKLYKVPDLDTTEYP 744
Db 61 FSKLYKVPDLDTMEYP 75

RESULT 7
Q8T069
ID Q8T069 PRELIMINARY; PRT; 620 AA.
AC Q8T069;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE LD23533P.
GN HUNTINGTIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY069515; AAL39660.1;
SQ SEQUENCE 620 AA; 69206 MW; 87F2240FD15C2432 CRC64;

Query Match 4.5%; Score 356.5; DB 5; Length 620;
Best Local Similarity 22.5%; Pred. No. 1.2e-13;
Matches 137; Conservative 116; Mismatches 268; Indels 89; Gaps 14;

```

```

QY 89 KELSATKKRVNCHLTICENIVASVRNSPEFQKLLGIAMLEFLCSDDAESDVRVMADE 148
Db 20 RNTECSOKOKITCFQOIAECIMSPSLAGHINYAAHCGTATNVLLFCEDVDVSRMSAE 79
QY 149 CLNKVITKALMDSNLPRIQLELYKEIKNGKNGAPRSALRALMFAELHLRQKCRPYLVNL 208
Db 80 NLANKILRSLEKTRVSRILMDYGEIKRNGNQRSLRCLNLFYVAPQIKHKKHVAVL 139
QY 209 LPLCTRTSKRPEESVQETLAAAVPKIMASPGNFANDNEIKVLKAFIANLKSSPTIRRT 268
Db 140 LQMTTISQRKTELLOETLQDFVYKHFSSRHTQQGLSDSECKLFFETLDQISSDCAVKRR 199
QY 269 AAGSAVSICQHSRRRTQVYFWSLLNVLLGLLPVVEDEHSTLLILGLVLLTLYLVPLQQQV 328
Db 200 SAQNCMSLIENARNRSLMARHGVNRYWELLTTOQQANS---VLGALGLLKLPLQIRGY 256
QY 329 KDTSLKSGFCVTRKEME---VSPSAEQLVQVYELTLH---HTQHDHNVVVTGALELQQL 382
Db 257 PGDSHEDSESLAGKQKQQOQTTSDCRQIIEYDYLHLLSTQHTANHAINA----- 309
QY 383 FRTPPPPELLQTLTAVGGIGQ-LTAAKEESGGRSRSISIVELIAGGSSSPVLSRKQ--- 438
Db 310 -----TLEVINGILQAVDAASDGGQCSQSLGSLRQLCLNQQLQHNEYLRKRSK 359
QY 439 -----KGKVLGEEALEDDSESRSDDVSSSALTASVKDEISGELAASSGVSTPFSAGHD 492
Db 360 NQIFQLKNYEVATSQHLEDENE-----DYDELVVGATAMQMKNSNAKLAQA 409
QY 493 IITEQPRSQHTLQADSVDLASCDLTSSATDGEDDEILSHSSSSQVSAVPDPAMDLDNGTQ 552
Db 410 KCREQQHQHQOQLE-VDNSLGINAG-----EDAPTEAPSSVADEGEP-----ESTK 457
QY 553 ASSPISDSSQTTTE-----GPDSSAVTPSSSIVLDGTDNQYLGLIQIGQPODEDEE 603
Db 458 LRCHIRNAKSISECVASDEKQGGHQRQORDEGIVVVAEDDDDD-----DDDDDDDD 510
QY 604 ATGILPDEASEAPRNSSMALQQAHLKKNMSHCHQPSDSSVDKFLVRDEATEPG---DQEN 660

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 13:58:08 ; Search time 85,9798 Seconds
(without alignments)
2391.326 Million cell updates/sec

Title: US-09-904-987-7
Perfect score: 7892
Sequence: 1 MATLEKLMKAFESLKSFOQQ.....DCIMASGRKSPQVRLCSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7821.5	99.1	3144	18 AAW09871	Human huntingtin.
2	7821.5	99.1	3144	20 AAY33493	Human huntingtin p
3	7821.5	99.1	3223	22 ABB11407	Human huntingtin's
4	7821.5	99.1	3223	22 ABB11470	Human huntingtin's
5	7805.5	98.9	3144	15 AAR58777	Protein encoded by
6	7805.5	98.9	3144	18 AAW36887	Previously undescri
7	7805.5	98.9	3144	19 AAW44742	Human huntingtin p
8	7145.5	90.5	3119	18 AAW36888	Mouse huntingtin p
9	7145.5	90.5	3119	18 AAW44743	Mouse huntingtin p
10	7041	89.2	3139	20 AAY08898	Human huntingtin's

11	2967.5	37.6	589	20 AAY33503	Human apopain clea
12	2784.5	35.3	552	20 AAY33502	Human apopain clea
13	2676.5	33.9	530	20 AAY33501	Human apopain clea
14	2589.5	32.8	513	20 AAY33500	Human huntingtin p
15	1974	25.0	418	18 AAW26545	Mouse huntingtin's
16	884.5	11.2	171	20 AAW90022	Human huntingtin p
17	674.5	8.5	3583	22 ABB4814	Drosophila melanog
18	456	5.8	108	20 AAW95071	Amino acid sequenc
19	456	5.8	108	20 AAW95076	Amino acid sequenc
20	416.5	5.3	79	22 AAB69616	Huntingtin's disea
21	213.5	2.7	86	20 AAW95073	GST-HD fusion prot
22	213.5	2.7	86	20 AAW95078	GST-HD fusion prot
23	210	2.7	94	20 AAW95075	GST-HD fusion prot
24	210	2.7	94	20 AAW95080	GST-HD fusion prot
25	210	2.7	192	22 AAB48292	GST-HD fusion prot
26	209	2.6	55	20 AAW95072	Human ZF4 protein.
27	209	2.6	55	20 AAW95077	GST-HD fusion prot
28	208	2.6	63	20 AAW95074	GST-HD fusion prot
29	208	2.6	63	20 AAW95079	GST-HD fusion prot
30	208	2.6	197	21 AAY83079	F-box protein FBP-
31	205	2.6	3536	22 ABB65480	Drosophila melanog
32	201.5	2.5	1954	22 ABB59197	Drosophila melanog
33	195.5	2.5	2783	13 AAR23962	AFP-1. Homo sapie
34	195.5	2.5	2783	13 AAR23963	AFP-1 (Ala 2460 Va
35	195.5	2.5	2783	22 AAB2946	Human androgen rec
36	194	2.5	99	22 AAO02076	Human polypeptide
37	193	2.4	85	22 AAO04573	Human polypeptide
38	193	2.4	995	22 ABB62510	Drosophila melanog
39	190	2.4	80	22 AAO04412	Human polypeptide
40	188	2.4	1368	22 AAB60262	Drosophila melanog
41	186.5	2.4	98	22 AAB69602	Huntingtin accumu
42	185.5	2.4	1151	22 ABB61598	Drosophila melanog
43	185	2.3	89	22 AAB69608	Huntingtin accumu
44	185	2.3	121	22 AAB69609	Huntingtin accumu
45	185	2.3	123	22 AAB69611	Huntingtin accumu

ALIGNMENTS

RESULT 1
AAW09871
ID AAW09871 standard; Protein; 3144 AA.
XX
AC AAW09871;
XX
DT 27-JUL-1997 (first entry)
XX
DE Human huntingtin.
XX
KW Huntingtin associated protein-1; HAP1; Huntington's disease.
XX
OS Homo sapiens.
XX
FH Key
FT Protein
FT Location/Qualifiers
FT 1..230
FT /note= "Claim 20"
XX
PN WO9717443-A1.
XX
PD 15-MAY-1997.
XX
PF 08-NOV-1996; 96WO-US17858.
XX
PR 09-NOV-1995; 95US-0556419.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Lanahan A, Li S, Li X, Ross CA, Sharp AH, Snyder S;
PI Worley PF;
XX
DR WPI; 1997-281032/25.
XX

PT Determination of the binding of huntingtin to huntingtin-associated
 PT protein-1 - useful for screening for drugs for treating or
 PT preventing Huntington's disease

XX Claim 20; Page 22-32; 69pp; English.

XX Human huntingtin (Hn) polypeptide (AAW09871) is the product of the
 CC Huntington's disease (HD) locus. Proteins which specifically bind
 CC to Hn, such as human huntingtin associated protein-1 (HAP1) (see
 CC also AAW09870), can be used in assays for screening drug candidates.
 CC The binding between Hn and HAP1 is enhanced by an expanded
 CC polyglutamine repeat in Hn, the length of which correlates with the
 CC time of disease onset. HAP1, in contrast to Hn, is expressed
 CC selectively in the brain, suggesting that it may contribute to the
 CC brain-specific pathology of HD. Hn, or portions of it, esp. amino
 CC acids 1-230, or yeast cells expressing Hn, can be used to identify
 CC cpds. that bind to, displace or prevent binding of Hn and HAP1.
 CC The method is useful for screening candidate drugs for treating,
 CC delaying onset of, or preventing HD.

XX Sequence 3144 AA;

Query Match 99.1%; Score 7821.5; DB 18; Length 3144;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
 DB 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 60
 QY 58 QPILLPQPP 117
 DB 61 QPILLPQPP 120
 QY 118 PEFQKLLGTAMEFLFICSDDAEDVRWVADCLNKVVKALMDSNLPRLQLELYKEIKNG 177
 DB 121 PEFQKLLGTAMEFLFICSDDAEDVRWVADCLNKVVKALMDSNLPRLQLELYKEIKNG 180
 QY 178 APRSLAALWRFALWRAHLVPOKRCRYLVNLLPCLTRTSKRPEESVQETLAAAPKIMAS 237
 DB 181 APRSLAALWRFALWRAHLVPOKRCRYLVNLLPCLTRTSKRPEESVQETLAAAPKIMAS 240
 QY 238 FGNFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSICQHSRRRTQYFYSWLLNVLLGL 297
 DB 241 FGNFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSICQHSRRRTQYFYSWLLNVLLGL 300
 QY 298 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVQDTSUKGSFGVTRKEMEYSPSAEQLVQVY 357
 DB 301 LVPVEDEHSTLLILGLVLLTLRYLVPLLQQQVQDTSUKGSFGVTRKEMEYSPSAEQLVQVY 360
 QY 358 ELTLHHTQHODHNWVTALELLOQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSRSG 417
 DB 361 ELTLHHTQHODHNWVTALELLOQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSRSG 420
 QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSDSDSVSSSALTASVKDEISGEL 477
 DB 421 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSDSDSVSSSALTASVKDEISGEL 480
 QY 478 AASSGVSTPGSAGHDITEOPRSQHTLQADSDVLDASCDLTSSATDGEDDILSHSSQVS 537
 DB 481 AASSGVSTPGSAGHDITEOPRSQHTLQADSDVLDASCDLTSSATDGEDDILSHSSQVS 540
 QY 538 AVPSDPAAMDNDGTQASSPISDSSQTTEGPDASVTPSDSSEIVLDGTDNQVLQIGQP 597
 DB 541 AVPSDPAAMDNDGTQASSPISDSSQTTEGPDASVTPSDSSEIVLDGTDNQVLQIGQP 600
 QY 598 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMCHCRQPSDSSVDKFLVRDEATEPGD 657
 DB 601 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMCHCRQPSDSSVDKFLVRDEATEPGD 660
 QY 658 QENKPCRIKIDIGQSTDDSDAPLVHCVRLLSASFLLTGTGKNVLPDPDRVRSVKALALSC 717
 DB 661 QENKPCRIKIDIGQSTDDSDAPLVHCVRLLSASFLLTGTGKNVLPDPDRVRSVKALALSC 720

QY 718 VGAVALHPSEFFSKLYKVPDITTEYEEQVSDILANYIDHGPQVRGATAILCGTLICS 777
 DB 721 VGAVALHPSEFFSKLYKVPDITTEYEEQVSDILANYIDHGPQVRGATAILCGTLICS 780
 QY 778 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTKLDESSVTCKLACTAVRNCVMSLC 837
 DB 781 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTKLDESSVTCKLACTAVRNCVMSLC 840
 QY 838 SSSYSELGLQIILDVLTNRSSYWLVTLELTAEDFRLVSFLEAKAENLRGAHHYT 897
 DB 841 SSSYSELGLQIILDVLTNRSSYWLVTLELTAEDFRLVSFLEAKAENLRGAHHYT 900
 QY 898 GLLKLQERVLNNVVIHLLGDPRVRHVAASLRLVKLFYKDCQDQADPVVAVARDQS 957
 DB 901 GLLKLQERVLNNVVIHLLGDPRVRHVAASLRLVKLFYKDCQDQADPVVAVARDQS 960
 QY 958 SVYKLLMHETQPPSHFSVTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1017
 DB 961 SVYKLLMHETQPPSHFSVTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1020
 QY 1018 RALTFGCEALCLLSTAFFPVCINSLGWHCVPPLSASDESKCTVGMATMILTLLSSAW 1077
 DB 1021 RALTFGCEALCLLSTAFFPVCINSLGWHCVPPLSASDESKCTVGMATMILTLLSSAW 1080
 QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKQEEVMPALGDRLVPMV 1137
 DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEANPAATKQEEVMPALGDRLVPMV 1140
 QY 1138 EQLFSHLLKVINICAHVLDVDPAPGPAIKALPSLTNPSPSLPIRRKKEKEPEQASVPL 1197
 DB 1141 EQLFSHLLKVINICAHVLDVDPAPGPAIKALPSLTNPSPSLPIRRKKEKEPEQASVPL 1200
 QY 1198 SPKKGSEASASRSQDTSQPVTTSSKSSLSGFSYHLPYSLKLDVLTATHANYKVTLDQN 1257
 DB 1201 SPKKGSEASASRSQDTSQPVTTSSKSSLSGFSYHLPYSLKLDVLTATHANYKVTLDQN 1260
 QY 1258 STEKFGGFLRSALDVLISOILELATLQDIDGKVEEILGYLKSFCSPREPMMATVCVQQLLK 1317
 DB 1261 STEKFGGFLRSALDVLISOILELATLQDIDGKVEEILGYLKSFCSPREPMMATVCVQQLLK 1320
 QY 1318 LFGNLAQFQGLSSNPSKSGRAQLGSSSRVPLGYHYCFMAYTHFTOALADASLRNM 1377
 DB 1321 LFGNLAQFQGLSSNPSKSGRAQLGSSSRVPLGYHYCFMAYTHFTOALADASLRNM 1380
 QY 1378 VQAEQENDTSGWFDVLOKYSTQLKNTLTSVTKNRADKNAIHNRILRLEPLVIKALKQYTT 1437
 DB 1381 VQAEQENDTSGWFDVLOKYSTQLKNTLTSVTKNRADKNAIHNRILRLEPLVIKALKQYTT 1440
 QY 1438 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDQVFTGFVLKQFEYIEVGFQFRESEAIIPNIF 1497
 DB 1441 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDQVFTGFVLKQFEYIEVGFQFRESEAIIPNIF 1500
 QY 1498 FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
 DB 1501 FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 2

AA33493

ID AA33493 standard; Protein; 3144 AA.

XX AC AA33493;

XX DT 19-JAN-2000 (first entry)

XX DE Human huntingtin polypeptide.

XX KW Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;

XX KW huntingtin polypeptide; Machado-Joseph disease; SCAL; SC2; SC6;

XX KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;

XX KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;

XX KW dentatorubropallidolysian atrophy; cell proliferation; cell survival;

Db 800 V G A A V A L H P E S F S K L Y K V P L D T T E P E Q Y V S D I L N Y I D H G D P Q V R G A T A I L C G T L I C S 859
Qy 778 I L S R S R F H V G D M W G T I R T L T G N T F S L A D C I P L L R K T L K D E S S V T C K L A C T A V R N C V M S L C 837
Db 860 I L S R S R F H V G D M W G T I R T L T G N T F S L A D C I P L R K T L K D E S S V T C K L A C T A V R N C V M S L C 919
Qy 838 S S S Y S E L G L Q L I D V L T L R N S S Y W L V R T E L L E T L A E I D F R L V S F L E A K A E N L H R C A H Y T 897
Db 920 S S S Y S E L G L Q L I D V L T L R N S S Y W L V R T E L L E T L A E I D F R L V S F L E A K A E N L H R C A H Y T 979
Qy 898 G L L K Q E R V L N V V I H L L G D E D P R V H V A A S L I R L V P K L F Y K C D Q G Q A D P V V A V A R D Q S 957
Db 980 G L L K Q E R V L N V V I H L L G D E D P R V H V A A S L I R L V P K L F Y K C D Q G Q A D P V V A V A R D Q S 1039
Qy 958 S Y L K L L M H E T O P P S H F S V T T I R Y R G Y N L L P S I T D V T M E N N L S R V A A V S H E L I T T 1017
Db 1040 S Y L K L L M H E T O P P S H F S V T T I R Y R G Y N L L P S I T D V T M E N N L S R V A A V S H E L I T T 1099
Qy 1018 R A L T F G C C E A L C L L S T A F P V C I W S L G W H C G V P P L S A S D E S R K S C T V G M A T M I L T L L S S A W 1077
Db 1100 R A L T F G C C E A L C L L S T A F P V C I W S L G W H C G V P P L S A S D E S R K S C T V G M A T M I L L L L S S A W 1159
Qy 1078 F P L D L S A H Q D A L I L A G N L L A S A P K S L R S S W A S E E A N P A A T K O B E V W P A L G D R A L V P M V 1137
Db 1160 F P L D L S A H Q D A L I L A G N L L A S A P K S L R S S W A S E E A N P A A T K O B E V W P A L G D R A L V P M V 1219
Qy 1138 E O L F S H L L K V I N I C A H V L D D V A P G P A I K A A L P S L T N P S L S P I R R K G K E P G E Q A S V P L 1197
Db 1220 E O L F S H L L K V I N I C A H V L D D V A P G P A I K A A L P S L T N P S L S P I R R K G K E P G E Q A S V P L 1279
Qy 1198 S P K G S E A S A A S R Q S D T S G P V T T S K S S L G S F Y H L P S Y L K L H D V L K A T H A N Y K V T L D L Q N 1257
Db 1280 S P K G S E A S A A S R Q S D T S G P V T T S K S S L G S F Y H L P S Y L K L H D V L K A T H A N Y K V T L D L Q N 1339
Qy 1258 S T E K E G G L R S A L D V L S Q I L E L A T Q D I G K C V E E I L G Y L K S C F S R E P M A T V C V Q Q L L K T 1317
Db 1340 S T E K E G G L R S A L D V L S Q I L E L A T Q D I G K C V E E I L G Y L K S C F S R E P M A T V C V Q Q L L K T 1399
Qy 1318 L F G T N L A S O F G L S N P S K S Q R A O L G S S V R P G L Y H Y C F M A P Y T H E T Q A L A D A S L R N M 1377
Db 1400 L F G T N L A S O F G L S N P S K S Q R A O L G S S V R P G L Y H Y C F M A P Y T H E T Q A L A D A S L R N M 1459
Qy 1378 V O A E N D T S G F D V L Q V S T Q L K T N L S V T K N R A D K N A I H N H I R L F E P L V I K A L Q Y T T 1437
Db 1460 V O A E N D T S G F D V L Q V S T Q L K T N L S V T K N R A D K N A I H N H I R L F E P L V I K A L Q Y T T 1519
Qy 1438 T T C V O L Q O V L D L L A Q L V O L R V N Y C L L D S D Q V F G F V L K Q E Y I E V G Q F R E S E A I I P N I F 1497
Db 1520 T T C V O L Q O V L D L L A Q L V O L R V N Y C L L D S D Q V F G F V L K Q E Y I E V G Q F R E S E A I I P N I F 1579
Qy 1498 F F V L L S Y E R V H S K Q I I G P K I I O L C D G I M A S G R K A 1533
Db 1580 F F V L L S Y E R V H S K Q I I G P K I I O L C D G I M A S G R K A 1615
RESULT 4
ABBI1470
ID ABBI1470 standard; peptide; 3223 AA.
XX AC ABBI1470;
XX DT 11-JAN-2002 (first entry)
XX DE Human Huntington's disease protein homologue, SEQ ID NO:1840.
XX KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibit; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;

bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antiulcer.
Homo sapiens.
WO200157188-A2.
09-AUG-2001.
05-FEB-2001; 2001WO-US03800.
03-FEB-2000; 2000US-0496914.
27-APR-2000; 2000US-0560875.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT;
WPI: 2001-457740/49.
N-PSDB; ABA08714.
Human proteins and DNA encoding sequences useful for preventing,
treating or ameliorating a medical condition in a mammalian subject
e.g. arthritis and cancer -
Claim 20; Page 195-196; 1963pp; English.
Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
haematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a novel human
polypeptide of the invention.
Sequence 3223 AA;
Query Match 99.1%; Score 7821.5; DB 22; Length 3223;
Best Local Similarity 99.8%; Pred. No. 0;

XX 05-MAR-1993; 93US-0027498.
 PR 01-JUL-1993; 93US-0085000.
 XX (GEO) GEN HOSPITAL CORP.
 PA Ambrose CM, Duyao WP, Gusella JF, MacDonald ME;
 PI WPI; 1997-558144/51.
 XX N-PSDB; AAT97924.
 DR Nucleic acid encoding huntingtin protein - useful for gene therapy
 PT of Huntington's disease
 XX Claim 2: Fig 4; 112pp: English.
 PS The present sequence represents a previously undescribed protein,
 CC encoded by a novel gene, termed huntingtin or htt. The huntingtin
 CC reading frame contains a polymorphic (CAG)_n trinucleotide repeat with at
 CC least 17 alleles in the normal population, varying from about 11 to 34
 CC CAG copies. Huntington's disease (HD) is a progressive neurodegenerative
 CC disorder characterised by motor disturbance, cognitive loss and
 CC psychiatric manifestations. The genetic defect causing HD is assigned to
 CC chromosome 4. On HD chromosomes, the length of the trinucleotide CAG
 CC repeat is substantially increased, e.g. about 37 to at least 73 copies.
 CC The huntingtin gene and proteins encoded by it, may be used for the
 CC diagnosis or treatment of Huntington's disease. The huntingtin gene
 CC is especially used in gene therapy of a symptomatic or presymptomatic
 CC patient. The method comprises providing a functional huntingtin gene with
 CC a (CAG)_n repeat of the normal range of 11-34 copies, or an antisense
 CC sequence, to the desired cells of the patient, in a manner that permits
 CC the expression of the huntingtin protein provided by the gene, or
 CC inhibits expression of the mutated huntingtin gene, for a time and in a
 CC quantity sufficient to provide the huntingtin function to the cells of
 CC the patient.
 XX Sequence 3144 AA:
 SQ

Query Match 98.9%; Score 7805.5; DB 18; Length 3144;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 57
 DB 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 60
 QY 58 QPLLPPQPP 117
 DB 61 QPLLPPQPP 120
 QY 118 PEFQKLGIAAMELFLICSDDAESDVRWVADECLNKVIKALMDSNLPRLQELYLEYKEIKNG 177
 DB 121 PEFQKLGIAAMELFLICSDDAESDVRWVADECLNKVIKALMDSNLPRLQELYLEYKEIKNG 180
 QY 178 APRSLRAALWRFELAEHLVRPQKRPYLVNLLPCLPRTSKRPESVQETLAAAVPKIMAS 237
 DB 181 APRSLRAALWRFELAEHLVRPQKRPYLVNLLPCLPRTSKRPESVQETLAAAVPKIMAS 240
 QY 238 FGNFANDNEIKVLKAFIAELKSSPTIRTAAGSAVSIQHSRRRTQYFYSWLLNVLGL 297
 DB 241 FGNFANDNEIKVLKAFIAELKSSPTIRTAAGSAVSIQHSRRRTQYFYSWLLNVLGL 300
 QY 298 LVPVEDHSTLLILGVLLTLRYLVPLLOQQVQKDTSLKSGFVTRKEMEVSPPSAEQLVQVY 357
 DB 301 LVPVEDHSTLLILGVLLTLRYLVPLLOQQVQKDTSLKSGFVTRKEMEVSPPSAEQLVQVY 360
 QY 358 ELTLHHTQHODHNVVGTGALVLLQQLFRTPPPELLQTLTAVGGTQTLTAKEESGGRSRG 417
 DB 361 ELTLHHTQHODHNVVGTGALVLLQQLFRTPPPELLQTLTAVGGTQTLTAKEESGGRSRG 420
 QY 418 SIVELIAGGGSSCPVLSRKQKGVLLGEEFALEDDESDSDVSSSALTATSVKDEISGEL 477
 DB 421 SIVELIAGGGSSCPVLSRKQKGVLLGEEFALEDDESDSDVSSSALTATSVKDEISGEL 480

QY 478 AASGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
 DB 481 AASGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
 QY 538 AVPSDPAWMDNDGTQASPSISDSQTTTEGPDSAVTSPDSSEIVLDGTNQYLGQIQGP 597
 DB 541 AVPSDPAWMDNDGTQASPSISDSQTTTEGPDSAVTSPDSSEIVLDGTNQYLGQIQGP 600
 QY 598 QDEDEEATGILPDEASEAFRNSMALQOAHLLKNMCHRCOPSSDSSVDKFLVLRDEATEPGD 657
 DB 601 QDEDEEATGILPDEASEAFRNSMALQOAHLLKNMCHRCOPSSDSSVDKFLVLRDEATEPGD 660
 QY 658 QENKPCRIRKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKGNVLPDRDRVSVKALALSC 717
 DB 661 QENKPCRIRKDGIGOSTDDSDAPLVHCVRLLSASFLLTGKGNVLPDRDRVSVKALALSC 720
 QY 718 VGAVALHPESFFSKLYKVPDLDTTEYPEEQVSDILNYIDHGDPOVRGATAILCGTILCS 777
 DB 721 VGAVALHPESFFSKLYKVPDLDTTEYPEEQVSDILNYIDHGDPOVRGATAILCGTILCS 780
 QY 778 ILSRSRHHVGMWGTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
 DB 781 ILSRSRHHVGMWGTIRLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
 QY 838 SSSYSELGLQLIIDLVLTRNSSYWLVRTELETLAEIDFLVSFLEAKAENLHRAHHYT 897
 DB 841 SSSYSELGLQLIIDLVLTRNSSYWLVRTELETLAEIDFLVSFLEAKAENLHRAHHYT 900
 QY 898 GLLKLQERVLNNVHLLGDEDPVRHVAAASLRLVPKLFYKCDQOQADPVAVARDQS 957
 DB 901 GLLKLQERVLNNVHLLGDEDPVRHVAAASLRLVPKLFYKCDQOQADPVAVARDQS 960
 QY 958 SVYLLKLMHETQPPSHFSVSITRIYRCYNLLPSITDVTMNNLSRVAASVSHELITSTT 1017
 DB 961 SVYLLKLMHETQPPSHFSVSITRIYRCYNLLPSITDVTMNNLSRVAASVSHELITSTT 1020
 QY 1018 RALTFGCCALCLLSTAFPVCISLWGHGCVPPPLSASDESRSKCTGVGMATMILLLSSAW 1077
 DB 1021 RALTFGCCALCLLSTAFPVCISLWGHGCVPPPLSASDESRSKCTGVGMATMILLLSSAW 1080
 QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLALVPMV 1137
 DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLALVPMV 1140
 QY 1138 EOLFSLHLKVINICAHVLDDVAPGPAIKAAPLSLTNPSPSLPIRRKKEKEPGEQASVPL 1197
 DB 1141 EOLFSLHLKVINICAHVLDDVAPGPAIKAAPLSLTNPSPSLPIRRKKEKEPGEQASVPL 1200
 QY 1198 SPKKGSEASASRSQSDTSGPVTTTSSKSLGFSFYHLPSYLKLDHVLKATHANYKVTLDLON 1257
 DB 1201 SPKKGSEASASRSQSDTSGPVTTTSSKSLGFSFYHLPSYLKLDHVLKATHANYKVTLDLON 1260
 QY 1258 STEKEGGFLRSALDVLISOILELATLQDQKGVVEETLGYLKCFSREPMMATVVCVQQLLKT 1317
 DB 1261 STEKEGGFLRSALDVLISOILELATLQDQKGVVEETLGYLKCFSREPMMATVVCVQQLLKT 1320
 QY 1318 LFGTNLASQFDGLSSNPSSKSGRAQRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
 DB 1321 LFGTNLASQFDGLSSNPSSKSGRAQRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1380
 QY 1378 VQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHILRFLPEPLVIKALKQYTT 1437
 DB 1381 VQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHILRFLPEPLVIKALKQYTT 1440
 QY 1438 TTCVQLOKQVLDLLAQLVQLRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEALIPNIF 1497
 DB 1441 TTCVQLOKQVLDLLAQLVQLRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEALIPNIF 1500
 QY 1498 FFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
 DB 1501 FFLVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 7

AAW44742
ID AAW44742 standard; Protein; 3144 AA.

AC AAW44742;

DT 01-JUN-1998 (first entry)

DE Human huntingtin protein.

KW Human; huntingtin gene; Huntington's disease; chromosome; marker;
KW locus; antisense; gene therapy; diagnosis.

OS Homo sapiens.

PN US5693757-A.

PD 02-DEC-1997.

PF 30-MAY-1995; 95US-0453265.

PR 20-MAY-1994; 94US-0246982.

PR 05-MAR-1993; 93US-0027498.

PR 01-JUL-1993; 93US-0085000.

PR 30-MAY-1995; 95US-0453265.

PA (GEO) GEN HOSPITAL CORP.

PI Ambrose CM, Duyao MP, Gusella JF, MacDonald ME;

XX WPI; 1998-031815/03.

DR N-PSDB; AAV05828.

XX Huntingtin protein and related nucleic acid - for diagnosis or
PT therapy of Huntington's disease

XX Claim 2; Fig 4; 112pp; English.

XX This is the amino acid sequence of the human huntingtin protein.
CC The gene sequence is characterised in that it contains a number of
CC CAG repeats in the 5' region (in this case 23 repeats). In healthy
CC individuals the usual number of CAG repeats in the gene is 11-34,
CC whereas in patients suffering from Huntington's disease (HD), the
CC number of repeats increases to 37-73 or 37-86. The huntingtin gene
CC spans 210 kb and encodes a protein of 348 kD. The gene is found in a
CC 500 kb region between the chromosomal markers D4S180 and D4S182 and is
CC preferentially mapped to the locus 4p16.3. The protein or the gene
CC encoding it, is useful for detecting a predisposition to develop HD,
CC for diagnosis and treatment of HD, especially by antisense and gene
CC therapy.

XX Sequence 3144 AA;

Query Match 98.98; Score 7805.5; DB 19; Length 3144;
Best Local Similarity 99.68; Pred. No. 0;
Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57

Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 60

Qy 58 QPLLPOQPP 117

Db 61 QPLLPOQPP 120

Qy 118 PEFKLLGIAMELFLCSDDAESDVRVVADECLNKV1KALMDSNLPRLQLELYKEIKNG 177

Db 121 PEFKLLGIAMELFLCSDDAESDVRVVADECLNKV1KALMDSNLPRLQLELYKEIKNG 180

Qy 178 APRSLRAALWRFALHVLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKTMA 237

Db 181 APRSLRAALWRFALHVLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKTMA 240

Qy 238 FGNFANDNEIKVLLKAFITANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFITANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 300
Qy 298 LVPVEDEHSTLLILGVLLTFLRYLVPLIQQQVKDTSLKGSFGVTRKEMEVSFSAEQLVQVY 357
Db 301 LVPVEDEHSTLLILGVLLTFLRYLVPLIQQQVKDTSLKGSFGVTRKEMEVSFSAEQLVQVY 360
Qy 358 ELTLHHTQHQNHNVTGALLELLOQLFRTPPELLOTLTAVGGIGQLTAKEESGGRSRG 417
Db 361 ELTLHHTQHQNHNVTGALLELLOQLFRTPPELLOTLTAVGGIGQLTAKEESGGRSRG 420
Qy 418 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVDEISGEL 477
Db 421 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVDEISGEL 480
Qy 478 AASGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASGVSTPGSAGHDIITEQPSOHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
Qy 538 AVSPDPAMDNDGTQASSPISDSQTTTEGPDPSAVTPSDSEIVLDGTDNQYLGLQIQGP 597
Db 541 AVSPDPAMDNDGTQASSPISDSQTTTEGPDPSAVTPSDSEIVLDGTDNQYLGLQIQGP 600
Qy 598 QDEDEEATGILPDEASEAFRNSMALQQAHLKMNHSCRPSPSSVDKFLVRDEATEPGD 657
Db 601 QDEDEEATGILPDEASEAFRNSMALQQAHLKMNHSCRPSPSSVDKFLVRDEATEPGD 660
Qy 658 QENKPCRIGDIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKALALSC 717
Db 661 QENKPCRIGDIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSKALALSC 720
Qy 718 VGAVALHPESFFSKLYKVPDLTTEPEEQYVSDILNYIDHGDPQVRGATAILCGTLICS 777
Db 721 VGAVALHPESFFSKLYKVPDLTTEPEEQYVSDILNYIDHGDPQVRGATAILCGTLICS 780
Qy 778 ILSRSRPHVGDWMTITRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 781 ILSRSRPHVGDWMTITRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
Qy 838 SSSYSELGLQIILDLVTLRNSYVWLVRTLETLAEIDRFLVSLFLEAKAENLRHGAHHT 897
Db 841 SSSYSELGLQIILDLVTLRNSYVWLVRTLETLAEIDRFLVSLFLEAKAENLRHGAHHT 900
Qy 898 GLKLEQVNLNVVLIHLGDEDEPRVRHVAASLIRLVPLKLFYKCDQCADPPVAVARDQS 957
Db 901 GLKLEQVNLNVVLIHLGDEDEPRVRHVAASLIRLVPLKLFYKCDQCADPPVAVARDQS 960
Qy 958 SVYKLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1017
Db 961 SVYKLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAVSHELITSTT 1020
Qy 1018 RALTFCCEALCLLSTAFVPCVWLSLGHGCVPLPLSASDESRSKCTVGMATMILTLLSSAW 1077
Db 1021 RALTFCCEALCLLSTAFVPCVWLSLGHGCVPLPLSASDESRSKCTVGMATMILTLLSSAW 1080
Qy 1078 FPLDLSAHQDALLIAGNLLAASAPKSLRSWSAEEANPAATKQEEVWPAIDRALVPVW 1137
Db 1081 FPLDLSAHQDALLIAGNLLAASAPKSLRSWSAEEANPAATKQEEVWPAIDRALVPVW 1140
Qy 1138 EQLFSHLLKVINICAHVLDVAVGPAIKAAPLSLTNPPSLSPIRRKKEKEPEQASVPL 1197
Db 1141 EQLFSHLLKVINICAHVLDVAVGPAIKAAPLSLTNPPSLSPIRRKKEKEPEQASVPL 1200
Qy 1198 SPKKGSEASAAARSQSDTSGPVTTSSKSSLSGSFYHLPSYLLKLDVLKATHANKVTLDLQN 1257
Db 1201 SPKKGSEASAAARSQSDTSGPVTTSSKSSLSGSFYHLPSYLLKLDVLKATHANKVTLDLQN 1260
Qy 1258 STEKFGGFLRSALDVLSQLLELATLQDIDGKCVEEILGYLKSCFSRPPMMATVVCVQOLKLT 1317
Db 1261 STEKFGGFLRSALDVLSQLLELATLQDIDGKCVEEILGYLKSCFSRPPMMATVVCVQOLKLT 1320

QY 960 YLKLHMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1019
 Db 941 YLKLHMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1000
 QY 1020 LTFGCEALCLLSTAPPVCWISLGHWCVPPLSASDESKKCTVGMATMILTLLSSAWFP 1079
 Db 1001 LTFGCEALCLLSTAPPVCWISLGHWCVPPLSASDESKKCTVGMATMILTLLSSAWFP 1060
 QY 1080 LDLSAHODALIAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPVMEQ 1139
 Db 1061 LDLSAHODALIAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPVMEQ 1120
 QY 1140 LFSHLLKVINICAHVLDVADPAIPAALPSLTNPSPSPSRKKEKEPEGEQASVPLSP 1199
 Db 1121 LFSHLLKVINICAHVLDVADPAIPAALPSLTNPSPSPSRKKEKEPEGEQASVPLSP 1180
 QY 1200 KKSASASASROSDTSGPVTTKSSSLGSFYHLPYLKLDVLAHANYKVTLDLQNST 1259
 Db 1181 KKSASASASROSDTSGPVTTKSSSLGSFYHLPYLKLDVLAHANYKVTLDLQNST 1240
 QY 1260 EKFGGFLRSALDVLSQLLELATLQDIGKCVVEILGYLKSCFSREPMATVCVQQLLTFL 1319
 Db 1241 EKFGGFLRSALDVLSQLLELATLQDIGKCVVEILGYLKSCFSREPMATVCVQQLLTFL 1300
 QY 1320 GYNLASQFDGLSSNPKSQGRLGSSSVRPLYHYCFMAYTHFTQALADASLRNMVQ 1379
 Db 1301 GYNLASQFDGLSSNPKSQGRLGSSSVRPLYHYCFMAYTHFTQALADASLRNMVQ 1360
 QY 1380 ABOENDTSGWEDVLQKVSQTLKNTLSTYTKRADKNATHNIRLFEPLVIAKQYTTT 1439
 Db 1361 ABOENDTSGWEDVLQKVSQTLKNTLSTYTKRADKNATHNIRLFEPLVIAKQYTTT 1420
 QY 1440 CVQLQKQVLDLLAQLVQLRVNYCLLDSQVFIGFVLKQFIEVYGOPRESEAIIPNIFFF 1499
 Db 1421 CVQLQKQVLDLLAQLVQLRVNYCLLDSQVFIGFVLKQFIEVYGOPRESEAIIPNIFFF 1480
 QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
 Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

RESULT 9

AAW44743

ID AAW44743 standard; Protein; 3119 AA.

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX AC AAW44743;

XX

PT

PT

XX

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches 1399; Conservative

46; Mismatches

68; Indels

21; Gaps

4;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

Huntingtin protein and related nucleic acid - for diagnosis or therapy of Huntington's disease

Example 9; Column 79-100; 112pp; English.

This is the amino acid sequence of the mouse huntingtin protein. The protein contains a Gln- and Pro-rich region (containing 12 and 27 Gln and Pro residues out of 42 residue respectively) in the N-terminal region. The gene sequence was isolated from a mouse PCC4 embryonal carcinoma cDNA library using the sequence of the human huntingtin gene (AAV05827) or probes and primers derived from the human huntingtin gene. The protein or the gene encoding it, is useful for detecting a predisposition to develop HD, for diagnosis and treatment of HD, especially by antisense and gene therapy.

Sequence 3119 AA;

Query Match

Best Local Similarity

Matches 1399; Conservative

46; Mismatches

68; Indels

21; Gaps

4;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY 960 YLKLHMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1019

Db 941 YLKLHMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1000

QY 1020 LTFGCEALCLLSTAPPVCWISLGHWCVPPLSASDESKKCTVGMATMILTLLSSAWFP 1079

Db 1001 LTFGCEALCLLSTAPPVCWISLGHWCVPPLSASDESKKCTVGMATMILTLLSSAWFP 1060

QY 1080 LDLSAHODALIAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPVMEQ 1139

Db 1061 LDLSAHODALIAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGDRALVPVMEQ 1120

QY 1140 LFSHLLKVINICAHVLDVADPAIPAALPSLTNPSPSPSRKKEKEPEGEQASVPLSP 1199

Db 1121 LFSHLLKVINICAHVLDVADPAIPAALPSLTNPSPSPSRKKEKEPEGEQASVPLSP 1180

QY 1200 KKSASASASROSDTSGPVTTKSSSLGSFYHLPYLKLDVLAHANYKVTLDLQNST 1259

Db 1181 KKSASASASROSDTSGPVTTKSSSLGSFYHLPYLKLDVLAHANYKVTLDLQNST 1240

QY 1260 EKFGGFLRSALDVLSQLLELATLQDIGKCVVEILGYLKSCFSREPMATVCVQQLLTFL 1319

Db 1241 EKFGGFLRSALDVLSQLLELATLQDIGKCVVEILGYLKSCFSREPMATVCVQQLLTFL 1300

QY 1320 GYNLASQFDGLSSNPKSQGRLGSSSVRPLYHYCFMAYTHFTQALADASLRNMVQ 1379

Db 1301 GYNLASQFDGLSSNPKSQGRLGSSSVRPLYHYCFMAYTHFTQALADASLRNMVQ 1360

QY 1380 ABOENDTSGWEDVLQKVSQTLKNTLSTYTKRADKNATHNIRLFEPLVIAKQYTTT 1439

Db 1361 ABOENDTSGWEDVLQKVSQTLKNTLSTYTKRADKNATHNIRLFEPLVIAKQYTTT 1420

QY 1440 CVQLQKQVLDLLAQLVQLRVNYCLLDSQVFIGFVLKQFIEVYGOPRESEAIIPNIFFF 1499

Db 1421 CVQLQKQVLDLLAQLVQLRVNYCLLDSQVFIGFVLKQFIEVYGOPRESEAIIPNIFFF 1480

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 1500 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533

Db 1481 LVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1514

QY 780 SRSRPHVGMCTIIRLTGNTFSLADCIPLLRKTKLDESSVTCKLACTAVRNCVMSLCSS 839
 DB 761 SRSRLVGEWGLNIRLTGNTFSLVDCIPLQKTLKDESSVTCKLACTAVRHCVLSLCSS 820
 QY 840 SYSELGLQIIDLTLRNSSYWLVRLTELETLAEIDFLVSLFLEAKAENLHRGAHHYGL 899
 DB 821 SYSDLGQLLIDMLPLKNSSYWLVRTELETLAEIDFLVSLFLEAKAENLHRGAHHYGF 880
 QY 900 LKQERVLNNVYHLGLDDBPRVHRVAASLRIRVPKLFYKCDQOQADPPVAVARDQSSV 959
 DB 881 LKQERVLNNVYHLGLDDBPRVHRVAASLRIRVPKLFYKCDQOQADPPVAVARDQSSV 940
 QY 960 YLKLWHETQPPSHFESVSTIRYGYNLSPITDVTWNNLSRVIAAVSHELITSTTRA 1019
 DB 941 YLKLWHETQPPSHFESVSTIRYGYNLSPITDVTWNNLSRVIAAVSHELITSTTRA 1000
 QY 1020 LTFGCCALCLLSTAFFVPCWISLGHGCGVPPPLSASDESRRKCTVGMATMILTLLSSAWFP 1079
 DB 1001 LTFGCCALCLLSTAFFVPCWISLGHGCGVPPPLSASDESRRKCTVGMATMILTLLSSAWFP 1060
 QY 1080 LOLSADQDALILAGNLLAASAPKSLRSSWASSEANPAATKQEEVWPALGDRLVPMVEQ 1139
 DB 1061 LOLSADQDALILAGNLLAASAPKSLRSSWASSEANPAATKQEEVWPALGDRLVPMVEQ 1120
 QY 1140 LFSHLLKVINICAHVLDVDPGPAIKALPSTNPPSLSPIRRKGEKEPGEQASVPLSP 1199
 DB 1121 LFSHLLKVINICAHVLDVDPGPAIKALPSTNPPSLSPIRRKGEKEPGEQASVPLSP 1180
 QY 1200 KKGSEASARSQSDTSGPVTTSSKSLGSFYHLPYSYLKLDHVLKATHANYKVTLDQNST 1259
 DB 1181 KKGSEASARSQSDTSGPVTTSSKSLGSFYHLPYSYLKLDHVLKATHANYKVTLDQNST 1240
 QY 1260 EKFGFLRSALDVLISOLELATLQDVGKVEEELGVLYKSCFREPMMATVCVQQLKTLF 1319
 DB 1241 EKFGFLRSALDVLISOLELATLQDVGKVEEELGVLYKSCFREPMMATVCVQQLKTLF 1300
 QY 1320 GTNLASQFGLSSNPSKSGRAQRGLSSVRPGLYHYCFMAYTHFTQALADASLRNMVQ 1379
 DB 1301 GTNLASQFGLSSNPSKSGRAQRGLSSVRPGLYHYCFMAYTHFTQALADASLRNMVQ 1360
 QY 1380 AEQDNTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNHRLFEPLVIKALKQYTTTT 1439
 DB 1361 AEQDNTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNHRLFEPLVIKALKQYTTTT 1420
 QY 1440 CVOLQKQVLDLQAQLVQLRVNYVCLLSDQVFGFVLKQFEYEVGQFRESEAIIPNIPFF 1499
 DB 1421 SVQLQKQVLDLQAQLVQLRVNYVCLLSDQVFGFVLKQFEYEVGQFRESEAIIPNIPFF 1480
 QY 1500 LVLLSYERYHSKQIIGIPKTIQICDGMASGRKA 1533
 DB 1481 LVLLSYERYHSKQIIGIPKTIQICDGMASGRKA 1514
 RESULT 10
 AAY08898
 ID AAY08898 standard; Protein; 3139 AA.
 XX
 AC AAY08898,
 XX
 DT 19-AUG-1999 (first entry)
 XX
 DE Human Huntington's chorea associated protein homologue.
 XX
 KW Huntington's chorea; human; model.
 XX
 OS Homo sapiens.
 XX
 PN JPI1137258-A.
 XX
 PD 25-MAY-1999.
 XX
 PF 14-NOV-1997; 97JP-0314020.

XX 14-NOV-1997; 97JP-0314020.
 PR (SLAK-) SLA KENKYUSHO KK.
 PA
 XX
 DR WPI; 1999-374378/32.
 DR N-PSDB; AAX78032.
 XX
 PT cDNA of huntington's gene - useful for development of animal
 model of Huntington's disease
 XX
 PS Claim 6; Page 10-16; 18pp; Japanese.
 XX
 CC This invention describes a novel human Huntington's chorea associated
 CC protein homologue. The cDNA and the other material of the invention
 CC are useful for the development of a model animal of Huntington's disease.
 CC
 XX SQ Sequence 3139 AA;
 Query Match 89.2%; Score 7041; DB 20; Length 3139;
 Best Local Similarity 88.9%; Pred. No. 0;
 Matches 1371; Conservative 70; Mismatches 85; Indels 16; Gaps 4;
 QY 1 MATLEKLMKAFESLKSTQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 60
 DB 1 MATLEKLMKAFESLKSTQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 55
 QY 61 LPQPP 120
 DB 56 --PP 113
 QY 121 QKLLGIAMELFLCDDAEDSVRMVADECLNKKVITKALMDSNLPRLQLELYKEIKKNGAPR 180
 DB 114 QKLLGIAMELFLCDDAEDSVRMVADECLNKKVITKALMDSNLPRLQLELYKEIKKNGAPR 173
 QY 181 SLRAALRRFAELAHVLRPKCRPYLVNLLPCLTSTRKPEESVQETLAAVPKIMASFGN 240
 DB 174 SLRAALRRFAELAHVLRPKCRPYLVNLLPCLTSTRKPEESVQETLAAVPKIMASFGN 233
 QY 241 FANDNEIKVLKKAFTANLKSSPTIRRTAGSAVSICQHSRRTOYFYSWLLNVLLGLLVP 300
 DB 234 FANDNEIKVLKKAFTANLKSSPTIRRTAGSAVSICQHSRRTOYFYSWLLNVLLGLLVP 293
 QY 301 VEDEHSTLLIIGVLLTLYRVLPLQQQVKDSLKSGFGVTRKENEVSPSAEQLVQVYELT 360
 DB 294 VEEHGTLLIIGVLLTLYRVLPLQQQVKDSLKSGFGVTRKENEVSPSTEQVQVYELT 353
 QY 361 LHHTOHQDHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGGRSGSIV 420
 DB 354 LHHTOHQDHNVVTGALELLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGGRSGSIV 413
 QY 421 ELIAGGSSCPVLSRKQKGVLLGEEAEDEDSERSDVSALTASVKDEISGELAAS 480
 DB 414 ELIAGGSSCPVLSRKQKGVLLGEEAEDEDSERSDVSALTASVKDEISGELAAS 473
 QY 481 SGVSTPGSA-----GHDIITEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 533
 DB 474 SGVSTPGSA-----GHDIITEQPRSOHTLQADSVDLASCDLTSSATDGEEDILSHS 533
 QY 534 SQVSAVPDPMALNDGTQASSPISDSSQTTTEGPDASVTPSDSSEIIVLDTQNYLGLQ 593
 DB 534 SQVSAVPDPMALNDGTQASSPISDSSQTTTEGPDASVTPSDSSEIIVLDTQNYLGLQ 593
 QY 594 IGOPO--DEDEATGILPDEASEAFNRNSMALQQAHLKKNMHSRQPSDSSVDKFLRDE 651
 DB 594 LGQPQDADADEAAGLLPHGGPDAPFRSSPTALQQLSHVLSLGHSRQPSDSSVDKFLRDE 653
 QY 652 ATEPGDQENKPCRIRKDGICOSTDDDSAPLVHCVRLLSASFLLTGGKNVLPDRDVRVSVK 711
 DB 654 AAEEAGEPENKPCRIRKDGICOSTDDDSAPLVHCVRLLSASFLLTGGKNVLPDRDVRVSVK 713
 QY 712 ALALSCVGAVALHPESFFSKLYKVPDLOTTTEYPEEQYVSDILNYIDHGDPOVRCATATLC 771

Db 714 ALALSCVGAVALHPSEFFSKLYKAPLDSVEYVDEQVSYIHGGDPQVRGATAILC 773
QY 772 GTLICSLSRSHRPHVGDWGTIHTLTGNTFSLADCTPLLRKTLKDESSVTCKLACTAVRN 831
Db 774 GTLVSLSRSHRPHVGDWGTIHTLTGNTFSLADCTPLLRKTLKDESSVTCKLACTAVRL 833
QY 832 CVMSLCSSSYSELGLOLIIDVLTNRSSYWLVRTELLETLAEIDFRLVFLFLEAKAENLHR 891
Db 834 CVAALCSCSSYVOWGLQITDILLTRSSYWLVRTELLETLAEIDFRLVFLFLEAKAENLHR 893
QY 892 GAHHYTGLLKQBRVNLNNVILHLLGDEDPDRHVAASLIRLVKFLYKCDQOGADPVVA 951
Db 894 GAHHYTGLLKQBRVNLNNVILHLLGDEDPDRHVAASLIRLVKFLYKCDQOGADPVVA 953
QY 952 VARDQSSVYLKLMHETQPPSHFSVSTIRYRGVNLPLSITDVTMNNLSRVIAAVSHE 1011
Db 954 VARDQSSVYLKLMHETQPPSHFSVSTIRYRGVNLPLSITDVTMNNLSRVIAAVSHE 1013
QY 1012 LITSTTRALTFCGCEALCLLSTAFVPCVWLSLGHVCGVPLGAADSRKSCVGMATTVLT 1071
Db 1014 LITSTTRALTFCGCEALCLLSTAFVPCVWLSLGHVCGVPLGAADSRKSCVGMATTVLT 1073
QY 1072 LLSAWFPLDLASHODALILAGNLLAASAPKSLRSSWASEEPAANPAATQEEVWPALGDR 1131
Db 1074 LLSAWFPLDLASHODALILAGNLLAASAPKSLRSSWASEEPAANPAATQEEVWPALGDR 1133
QY 1132 ALVPVVEQLFSLHLLKVINICAHVLDVDPGPAKALPSLTNPSPSLPTRRKGKEPGE 1191
Db 1134 ALVPVVEQLFSLHLLKVINICAHVLDVDPGPAKALPSLTNPSPSLPTRRKGKEPGE 1193
QY 1192 QASVPLSPKKGSEASARSQSDTSGPTVTSKSSLSGFYHLPYSLKLDHVLKATHANKV 1251
Db 1194 QASVPLSPKKGSEASARSQSDTSGPTVTSKSSLSGFYHLPYSLKLDHVLKATHANKV 1253
QY 1252 TLDLQNSTEKGFEGLRSALDVLQLELATLQDICKCYVEEILGYLKSCFSREPMMATVCV 1311
Db 1254 TLDLQNSTEKGFEGLRSALDVLQLELATLQDICKCYVEEILGYLKSCFSREPMMATVCV 1313
QY 1312 QOLLKTLFTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAYTHFTQALAD 1371
Db 1314 QOLLKTLFTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAYTHFTQALAD 1373
QY 1372 ASLRNVAQENDSRGFDVLQKYSTQIKNLTSTVTKNRADKNAIHNRILFEPPLVKA 1431
Db 1374 ASLRNVAQENDSRGFDVLQKYSTQIKNLTSTVTKNRADKNAIHNRILFEPPLVKA 1433
QY 1432 LKQYTTTTVCVQKQVLDLLAQLVQLRVNYCLLSDQVFIQFVLKQFYEIVGOFRESEA 1491
Db 1434 LKQYTTTTVCVQKQVLDLLAQLVQLRVNYCLLSDQVFIQFVLKQFYEIVGOFRESEA 1493
QY 1492 IIPNIFPFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1494 IIPNIFPFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1535

RESULT 11
AAV33503
ID AAV33503 standard; Protein; 589 AA.
XX
AC AAV33503;
XX
DT 19-JAN-2000 (first entry)
XX
DE Human apopain cleavage fragment 3.
XX
KW proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW huntingtin polypeptide; Machado-Joseph disease; SCAL; SCAG6;
KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW dentatorubropallidolysian atrophy; cell proliferation; cell survival;
KW neoplastic; malignant; autoimmune; fibrotic; apopain.
OS Homo sapiens.

XX WO9945944-A1.
PN 16-SEP-1999.
XX
PD 11-MAR-1999; 99WO-US05250.
PF 12-MAR-1998; 98US-0041886.
PR (BURN-) BURNHAM INST.
XX Bredesen DE, Rabizadeh S;
PI WPI; 1999-561617/47.
XX
DR New proapoptotic dependence peptides, used to develop products for
PT treating, e.g. Alzheimer's disease -
PS Disclosure; Page 176-178; 1999p; English.
XX
CC This invention describes novel pure proapoptotic dependence peptides
CC which comprise a sequence of an active dependence domain selected from
CC dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC huntingtin polypeptide, Machado-Joseph disease gene product, SCAL, SCAG2,
CC SCAG6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC of inducing cell death and can be used to develop products to mediate or
CC inhibit apoptosis. The methods can be used for reducing the severity of
CC a proapoptotic dependence domain mediated pathological conditions e.g.
CC Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC Spinocerebellar ataxia, dentatorubropallidolysian atrophy,
CC Machado-Joseph disease, stroke or head trauma. They can also be used for
CC reducing the severity of a pathological condition mediated by upregulated
CC cell proliferation or cell survival e.g. neoplastic, malignant,
CC autoimmune or fibrotic conditions. This sequence represents a human
CC apopain cleavage fragment described in the method of the invention.
XX
SQ Sequence 589 AA;

Query Match 37.68; Score 2967.5; DB 20; Length 589;
Best Local Similarity 99.5%; Pred. No. 4.2e-205;
Matches 586; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MATLEKLMKAFESLKST---QQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSTFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 60
QY 58 QPLLPPQPP 117
Db 61 QPLLPPQPP 120
QY 118 PEFQKLLGIAMELFLLCSDDAESDVRVMADECLNKVKIKALMDSNLPRLQLELYKEIKKNG 177
Db 121 PEFQKLLGIAMELFLLCSDDAESDVRVMADECLNKVKIKALMDSNLPRLQLELYKEIKKNG 180
QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLPRTSKRPESVQETLAAAVPKTMS 237
Db 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLPRTSKRPESVQETLAAAVPKTMS 240
QY 238 FGNFANDNETKVLKAFIANLKSSSPPIRTTAAGSAVSICQHSRRTOYFYFWLLNVLGL 237
Db 241 FGNFANDNETKVLKAFIANLKSSSPPIRTTAAGSAVSICQHSRRTOYFYFWLLNVLGL 300
QY 298 LVPVEDHSTLLTILGVLLTLRYLPLLOQVQKDTSLKSGFVTRKEMESVPSAEQLVQVY 357
Db 301 LVPVEDHSTLLTILGVLLTLRYLPLLOQVQKDTSLKSGFVTRKEMESVPSAEQLVQVY 360
QY 358 ELTLHHTQHODHNWVTALELLQQLFRTPELQTLTAVGGIGQLTAAKEESGGRSRG 417
Db 361 ELTLHHTQHODHNWVTALELLQQLFRTPELQTLTAVGGIGQLTAAKEESGGRSRG 420
QY 418 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRSDVSSSALTASVKDEISGEL 477
Db 421 SIVELIAGGSSCSPVLSRKQKGVLLGEEALEDDSESRSDVSSSALTASVKDEISGEL 480

XX PS Disclosure; Page 173-174; 199pp; English.

XX CC This invention describes novel pure proapoptotic dependence peptides

CC CC which comprise a sequence of an active dependence domain selected from

CC CC dependence polypeptides consisting of p75NTR, androgen receptor, DCC,

CC CC huntingtin polypeptide, Machado-Joseph disease gene product, SCAl, SCa2,

CC CC SCa6 and atrophin-1 polypeptide. The proapoptotic peptides are capable

CC CC of inducing cell death and can be used to develop products to mediate or

CC CC inhibit apoptosis. The methods can be used for reducing the severity of

CC CC a proapoptotic dependence domain mediated pathological conditions e.g.

CC CC Huntington's disease, Alzheimer's disease, Kennedy's disease,

CC CC Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,

CC CC Machado-Joseph disease, stroke or head trauma. They can also be used for

CC CC reducing the severity of a pathological condition mediated by upregulated

CC CC cell proliferation or cell survival e.g. neoplastic, malignant,

CC CC autoimmune or fibrotic conditions. This sequence represents a human

CC CC apocain cleavage fragment described in the method of the invention.

XX SQ Sequence 530 AA;

Query Match 33.9%; Score 2676.5; DB 20; Length 530;
 Best Local Similarity 99.4%; Pred. No. 3.5e-184;
 Matches 527; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57

Db 1 MATLEKLMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 60

QY 58 QPLLPOPPQA 57

Db 61 QPLLPOPPQA 60

QY 118 PEFQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKKNG 177

Db 121 PEFQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKKNG 180

QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237

Db 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240

QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 297

Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 300

QY 298 LVPVEDEHSTLLILGVLTLRLYLVPLLQQQVKDTSLSKGSFGVTRKEMEVSFSAEQLVQVY 357

Db 301 LVPVEDEHSTLLILGVLTLRLYLVPLLQQQVKDTSLSKGSFGVTRKEMEVSFSAEQLVQVY 360

QY 358 ELTLHHTQHODHNVTGALQLQLPRTPPPELLQTLTAVGGIGQITAAKESGGRSRSG 417

Db 361 ELTLHHTQHODHNVTGALQLQLPRTPPPELLQTLTAVGGIGQITAAKESGGRSRSG 420

QY 418 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSDSRSDVSSSALTASVKDEISGEL 477

Db 421 SIVELIAGGSGSPVLSRKQKGVLLGEEALEDDSDSRSDVSSSALTASVKDEISGEL 480

QY 478 AASSGVSTPGSAGHDIIITEQPSRQHTLQADSVDLASCDLTSSATDGDDED 527

Db 481 AASSGVSTPGSAGHDIIITEQPSRQHTLQADSVDLASCDLTSSATDGDDED 530

RESULT 14

AY33500
 ID AAY33500 standard; Protein; 513 AA.

XX AC AAY33500;

XX DT 19-JAN-2000 (first entry)

XX DE Human huntingtin polypeptide amino terminal fragment.

XX KW Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;

KW huntingtin polypeptide; Machado-Joseph disease; SCAl; SCa2; SCa6;
 KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
 KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
 KW dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
 KW neoplastic; malignant; autoimmune; fibrotic.

OS Homo sapiens.

PN WO9945944-A1.

XX PD 16-SEP-1999.

XX PF 11-MAR-1999; 99WO-US05250.

XX PR 12-MAR-1998; 98US-0041886.

XX PA (BURN-) BURNHAM INST.

XX PI Bredesen DE, Rabizadeh S;

XX DR WPI; 1999-561617/47.

XX PT New proapoptotic dependence peptides, used to develop products for
 treating, e.g. Alzheimer's disease -

XX XX Disclosure; Page 171-172; 199pp; English.

This invention describes novel pure proapoptotic dependence peptides which comprise a sequence of an active dependence domain selected from dependence polypeptides consisting of p75NTR, androgen receptor, DCC, huntingtin polypeptide, Machado-Joseph disease gene product, SCAl, SCa2, SCa6 and atrophin-1 polypeptide. The proapoptotic peptides are capable of inducing cell death and can be used to develop products to mediate or inhibit apoptosis. The methods can be used for reducing the severity of a proapoptotic dependence domain mediated pathological conditions e.g. Huntington's disease, Alzheimer's disease, Kennedy's disease, Spinocerebellar ataxias, dentatorubropallidoluysian atrophy, Machado-Joseph disease, stroke or head trauma. They can also be used for reducing the severity of a pathological condition mediated by upregulated cell proliferation or cell survival e.g. neoplastic, malignant, autoimmune or fibrotic conditions. This sequence represents the human huntingtin polypeptide amino terminal fragment described in the method of the invention.

SQ Sequence 513 AA;

Query Match 32.8%; Score 2589.5; DB 20; Length 513;
 Best Local Similarity 99.4%; Pred. No. 6.2e-178;
 Matches 510; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 57

Db 1 MATLEKLMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 60

QY 58 QPLLPOPPQA 57

Db 61 QPLLPOPPQA 60

QY 118 PEFQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKKNG 177

Db 121 PEFQKLLGAMELFLCSDDAESDVRVADECLNKVKIKALMDSNLPRLQLEYKEIKKNG 180

QY 178 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237

Db 181 APRSLRAALWRFALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240

QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 297

Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTRRTAAGSAVSTICQHSRRTQYFYSWLLNVLLGL 300

QY 298 LVPVEDEHSTLLILGVLTLRLYLVPLLQQQVKDTSLSKGSFGVTRKEMEVSFSAEQLVQVY 357

Db 301 LVPVEDEHSTLLILGVLTLRLYLVPLLQQQVKDTSLSKGSFGVTRKEMEVSFSAEQLVQVY 360


```
QY 241 FANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSTCOHSRSTQYFYSWLLNVLLGLLVP 300
Db 241 FANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSTCOHSRSTQYFYSWLLNVLLGLLVP 300
QY 301 VEDEHSTLLILGLVLLTRVLPVLLQOVKDTSLKGSFGVTRKEMEVSPSAEQLVQVYELT 360
Db 301 VEDEHSTLLILGLVLLTRVLPVLLQOVKDTSLKGSFGVTRKEMEVSPSAEQLVQVYELT 360
QY 361 LHHTQHQNHNVTGALVLELLOQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSGSIV 420
Db 361 LHHTQHQNHNVTGALVLELLOQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSGSIV 420
QY 421 ELTAGGSCSPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVKDRISGELAAS 480
Db 421 ELTAGGSCSPVLSRKQKGVLLGEEALEDDSESRSVSSSALTASVKDRISGELAAS 480
QY 481 SGYSTPGSAGHDITTEQPRSOHTLOADSVLDLASCIDLTSATDGEDILSHSSSQSVAVP 540
Db 481 SGYSTPGSAGHDITTEQPRSOHTLOADSVLDLASCIDLTSATDGEDILSHSSSQSVAVP 540
QY 541 SDPAMLDNGTQASSPISDSSTTTTGGPSAVTPSDSSEIVLDGTDNQVLGILQIQPQDE 600
Db 541 SDPAMLDNGTQASSPISDSSTTTTGGPSAVTPSDSSEIVLDGTDNQVLGILQIQPQDE 600
QY 601 DEBATGLPDEASEAFRNSMALQQAHLKNNKSHCHRCQPSDSSVDKFLVLRDEATEPGDQEN 660
Db 601 DEBATGLPDEASEAFRNSMALQQAHLKNNKSHCHRCQPSDSSVDKFLVLRDEATEPGDQEN 660
QY 661 KPCRIGKIDIGQSTDDPSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALASCVGA 720
Db 661 KPCRIGKIDIGQSTDDPSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALASCVGA 720
QY 721 AVALHPESFSKLYKVPDLTTEPEQYVSDIILNYIDHGDQVQRGATAILCGTLICSIILS 780
Db 721 AVALHPESFSKLYKVPDLTTEPEQYVSDIILNYIDHGDQVQRGATAILCGTLICSIILS 780
QY 781 RSRFHVGDMWGTITRTLTGNFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
Db 781 RSRFHVGDMWGTITRTLTGNFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLCSSS 840
QY 841 YSELGLLIIIDVLTNRSSVWLVRTLELTLAEIDFLVSLFLEAKAENLHRGAHHYTGLL 900
Db 841 YSELGLLIIIDVLTNRSSVWLVRTLELTLAEIDFLVSLFLEAKAENLHRGAHHYTGLL 900
QY 901 KLQERVLNNVVIHLGDEDPVRHVHAAASLIRLVPLFKYKCDQOGADPVVAVARDQSSVY 960
Db 901 KLQERVLNNVVIHLGDEDPVRHVHAAASLIRLVPLFKYKCDQOGADPVVAVARDQSSVY 960
QY 961 LKLLMHETQPPSHFSVTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRAL 1020
Db 961 LKLLMHETQPPSHFSVTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRAL 1020
QY 1021 TFGCCEALCLLSTAFFVCIVNSLGHWCVPPLSASDESRSKCTVGMATMILTLLSSAWFPL 1080
Db 1021 TFGCCEALCLLSTAFFVCIVNSLGHWCVPPLSASDESRSKCTVGMATMILTLLSSAWFPL 1080
QY 1081 DLSAHQDALILAGNLLAASAPKSLRSSWASSEEPANPAATKQEEVWPALGDRALVPVWEOL 1140
Db 1081 DLSAHQDALILAGNLLAASAPKSLRSSWASSEEPANPAATKQEEVWPALGDRALVPVWEOL 1140
QY 1141 FSHLLKVINICAHVLDVAVCPAIPAALPSITNPPSLPIRRKGKEKEPGEQASVPLSPK 1200
Db 1141 FSHLLKVINICAHVLDVAVCPAIPAALPSITNPPSLPIRRKGKEKEPGEQASVPLSPK 1200
QY 1201 KGSASAASROSQDSGPVTTSSKSSLSGFYHPLSYLKLHDLVLAATHANYKVTLDLQNSTE 1260
Db 1201 KGSASAASROSQDSGPVTTSSKSSLSGFYHPLSYLKLHDLVLAATHANYKVTLDLQNSTE 1260
QY 1261 KFGGFLRSALDVLSQLLELATLQDIDGKVEEILGYLKSFSRPPMATVVCQQLLTLFG 1320
Db 1261 KFGGFLRSALDVLSQLLELATLQDIDGKVEEILGYLKSFSRPPMATVVCQQLLTLFG 1320
```

```
QY 1321 TNLASQFDGLSSNPSKSGQRAQLGSSSVRGLYHYCFMAYPYTHFTQALADASLRNNVQA 1380
Db 1321 TNLASQFDGLSSNPSKSGQRAQLGSSSVRGLYHYCFMAYPYTHFTQALADASLRNNVQA 1380
QY 1381 BOENDTSGWFDVLQKVSTQLKTNLTSTVKKNRADKNAIHNHRLFEPLVIALKQYTTTTC 1440
Db 1381 BOENDTSGWFDVLQKVSTQLKTNLTSTVKKNRADKNAIHNHRLFEPLVIALKQYTTTTC 1440
QY 1441 VOLQKQVLDLQAQLVQLRVNVCLLDSDQVFGFVVKQFEVIEVQGFRESEAIIPNIFFL 1500
Db 1441 VOLQKQVLDLQAQLVQLRVNVCLLDSDQVFGFVVKQFEVIEVQGFRESEAIIPNIFFL 1500
QY 1501 VLLSYERYHSHKIIIGIPKIIQIOLCDGIMASGRKASQPYPRLCSP 1543
Db 1501 VLLSYERYHSHKIIIGIPKIIQIOLCDGIMASGRKASQPYPRLCSP 1543
```

RESULT 2

```
US-10-042-417-32
; Sequence 32, Application US/10042417
; Patent No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042.417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-32
```

```
Query Match 2.6%; Score 208; DB 12: Length 197;
Best Local Similarity 32.3%; Pred. No. 9.3e-06;
Matches 63; Conservative 12; Mismatches 60; Indels 60; Gaps 7;
QY 27 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 77
Db 6 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 62
QY 78 -----AAVEE-----PLHR-----PKKELSATKKDVRNHNCLTICENI--VAQSV 114
Db 63 DDDVPADWVAEESGPGAQNSPYQLRRKTLTPKRTACTPKNSMEGASTSTTFNFGHRAKRA 122
QY 115 RNSPEFQKLLGIAMBELFLCSDDAESDVRMVADECLNKVIRKALMDSNL----- 162
Db 123 RVSGKSQDLSAAPAEQYL-----OEKLPDEVVLKIFSYLLRQDLCLRAACVCKRESE 173
QY 163 ----PRQLQELYKEI 173
Db 174 LANDPNLMKRLYMEV 188
```

RESULT 3

```
US-09-770-689A-5
; Sequence 5, Application US/09770689A
; Patent No. US2002011517A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001079
; CURRENT APPLICATION NUMBER: US/09/770,689A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
```


QY	154	IKALMDSNPRLOLELYKEIKKNGAPRSRAALAHVLRPKQCRPYLVNLLPCIT	213	
Db	233	-----ASUPSL-----PLSAASFRKE-----YLGNSLSTVLP	260	
QY	214	RTSKREESVOETLAAAVPKIMASGFNFANDNEIKVLLKAFIANLKSS---	SPTIRRTAA	270
Db	261	-----TEGTLOQNVSEASKEKATLLIDRLDTEFSELEYSEMGSFVSP----	KA	310
QY	271	GSAVSTCOHSRRTQYFYSWLLNVLLGLLVPVDEHSTLLILGVLLTLRLVLP	LOQOVKD	330
Db	311	ESAVIVA-----NPREIIVYKNKDEEKLVSNNILHNQOELPTALTKLVE	356	
QY	331	TSL-----KGSFGVTRKMEVSPSAEQLVQVVELTLHHTQHDHNVVTGA	LELLQQLFR	384
Db	357	DEVVSEKAKDSNEKRVAE-APMREEVADP-----	387	
QY	385	TPPELILQTLTAVGGICQLTAAKEESGGRSRGSGISVELIAGGSSCPVLS	RKQKGVLL	444
Db	388	-KPE-----RWVEYKDSKEDS-----DMLAAGKTESNLESVKDK-K	CEFA	426
QY	445	G--EEEALEDSDSRSDVSSALT-ASVKDETSGEL-----AASSGVST-	P	486
Db	427	DSLEQTNHEKDSSESSNDTSPSTPEGIKDRSGAYITCAPFNPAAPESIA	TATIFPLLGP	486
QY	487	GSAGH-----DIITEQPRSOHT-----LOADSVDLASCDLTSSA	520	
Db	487	TSENKTDKKEIKBAQIVTEKNTSTKTNPFLVAQADSETDYVTTDNLTK	VYEEVANN	546
QY	521	TGDEEDILSHS--SSOVSAPVS-----DPAMDLDNGTQASSPIDS-	-----SOTTTEGPDS	570
Db	547	PEGULTDVOGACESELNEVTGKIAIYTKMDL--VQTVSEVMQESLYPAA	QCLPSFEES	603
QY	571	AVTPSD--SSBIVLDGTDNOYL--GLQIGOPODEDEATGI-----	607	
Db	604	EATPSPVLPIVMEAPLNSAVPSAGASVQSPSSPLEASSVNVYESIKHEP	NPVPPEEAM	663
QY	608	-----LPDEASAFNRSSMALQAH-----LLKNMSHCRQSDSSVDK	FVL-	648
Db	664	SVSLKKVSGIKEEKP-ENINAALOTEAPYISACDLTKLSAEPAPDFS	YSEMA	722
QY	649	-----RDEATEPGDOENKPCRIGDIDGSTDSDSAPLV-----	HCVRLLLSASFLLTG	695
Db	723	KVEQVPDHSSELVEDSDSPDEVDL-----FSDSIDPVPQKQDET	WMLVKESTETS	775
QY	696	GKNVL-----VPDROVRVSVKALALSCVGAVALHPE--SFFSKLYK	PLDPTTEY	743
Db	776	FESMIEYENKEKLSALPPEGGKPYLSFKLSLSDNTKDTLLPDEVSTLS	KKKEKPIQMEEL	835
QY	744	PEEQVSDIL	753	
Db	836	STAVYSNDL	845	

RESULT 6

RESOLUTION 6
US-09-972-599A-6
Sequence 6, Application US/09972599A
Patent No. US20020077295A1
GENERAL INFORMATION:
APPLICANT: STRITTMATTER, STEPHEN M.
TITLE OF INVENTION: NOGO-RECEPTOR-MEDIATED BLOCKADE OF AXONAL GROWTH
FILE REFERENCE: C077 CIP US
CURRENT APPLICATION NUMBER: US/09/972,599A
CURRENT FILING DATE: 2001-10-06
PRIOR APPLICATION NUMBER: ECT/US01/01041
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 09/758,140
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/236,378
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/207,366
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/175,707

RESULT 7

Qy	608	-----LPDEASAFRNSMALQQA-----LLKNMSHCROPSSDSSVKFVL-	648
Db	664	SVSLKKVSGTKEELKEP-ENINAALQETAPYISACDLIKETKLSAEPDFSDYSEMA	722
Qy	649	-----RDEATEPGDQENKPKRIKGDICQSDDDSDAPLV-----HCVRLLSASFLLTG	695
Db	723	KVEQVPVDHSELVEDSSPDSEPYDL-----FSDDSIPDVQKQDETVMVKESLTETS	775
Qy	696	GKNVL-----VPDRVRVSVKALALSCVGAVALHPE--SFFSKLYKVPLOTTEY	743
Db	776	FESMIEYENKEKLSALPPEGGKPYLSFSLSDNTKDTLLPDEVSTLSKKKXIPLOMUEL	835
Qy	744	PEEQVSDIL	753
Db	836	STAVYSNDL	845

RESULT 8

US-09-893-348-23
 ; Sequence 23, Application US/09893348
 ; Patent No. US20020072493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EISENBACH-SCHWARTZ, Michal
 ; APPLICANT: COHEN, Irun R.
 ; APPLICANT: BESERMAN, Pierre
 ; APPLICANT: MOSONOGO, Alon
 ; APPLICANT: MOALEM, Gila
 ; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR

Query Match 2.3%; Score 178.5; DB 10; Length 1192;
Best Local Similarity 19.4%; Pred. No. 0.0067;
Matches 165; Conservative 114; Mismatches 302; Indels 269; Gaps 36;

Qy	34	QOQQPPPPPPPPPPQLPQPQQAPLQPLPQQPPPPPPPPPPGPAVAEBPLHRPKKELSA	93
Dd		: : : : : : : :	
Dd	135	EDEEPARPPPPPPASV---SPQAEVPWTPPAPAAPPSTPA-----APRRGSS	182
Qy	94	TKKDRVNCLTICENIVAQSVRNSPEQKLGLGTAMELFLLCSDDAESDVRYVADCLNKV	153
Dd		: : : : : : : : : : :	
Dd	183	GVSDETLPALPAASEPVIRSSAENNMDLKEOPGNITSA-----GOEDFPSSLLETA---	232
Qy	154	IKALMDSNLRLQLELYKEIKKGAPSRSLRAALWRFAELAHLRPOKCRPYLVNLLPCLT	213
Dd		: : : : : : : : : : : :	
Dd	233	-----ASLPSLS-----PLSAAFSKEHE-----YLGNLSTVL P	260
Qy	214	RTSKRPEBSVOETLAAPVKIMASFGFANFANDNEIKULLKAFTIANLKSS---SPTIRRTAA	270
Dd		: : : : : : : : : : :	
Dd	261	-----TEGTLENSEAKSEYSEKAKTLLDRODTFESELEYSEMGSFSVSP-----KA	310
Qy	271	GSAVICQHRSRTQVFYSWLLNVLLGLLVPEDEHSTLLILGVLLTRLVLPVLOQQVKD	330
Dd		: : : : : : : : : :	
Dd	311	ESAVIVA-----NPREEIIVKNKDBEEKLYSNLILHNQOQLPATLTKLYKE	356
Qy	331	TSL-----KSGFVGTRKEMVEVSPSABOLVOVYVELTLHHQTQHQDHNVVTGALELLQQLFR	384
Dd		: : : : : : : : : : : :	
Dd	357	DEWSSERKAKDSFKRVAVE-APMBREYADF-----DMLAAGKKTESLNKVDK-KCFA	426
Qy	385	TPPPLLTLTAGVIGIGOLTAARESGGRSGSI VIELIAGGSGSCSPVLSRKQKCKVLL	444
Dd		: : : : : : : : : : : :	
Dd	388	KPFE-----RVMEVDKSREDS-----AAASSGVST-----P	486
Qy	445	G---EEEALEDDESRSDDSVSSALT-ASVKDEISGEL-----SOTTITEGPDS	570
Dd		: : : : : : : : : : :	
Dd	427	DSLQTNHKESDNDDTSPSTPEGIKORPCAYITCAPFNPAATESIATNIPFLGGDP	486
Qy	487	GSAGH-----DLITFOPRSQHT-----LOADSVDLASCOLDTSSA	520
Dd		: : : : : : : : : : :	
Dd	487	TSENKTDKKLT EEKQAIVTEKNTSTKTSNPFVLAQDSETDYVTVDNLTUKYTEEVANM	546
Qy	521	TGDGEDILSHS-SQOVSAPVS-----DPAMDNDGTQASSPISDS-----SOTTTTEGPDS	570
Dd		: : : : : : : : : : : :	
Dd	547	PEGLTPDLVQEACESELNEVTGTKIAVETKM DL--VQTSVMQESLYPAAQLCSFFES	603
Qy	571	AVTSPD-SSEIVLGDCTNOYL---GLQIQOPQODEEATGI-----	607
Dd		: : : : : : : : : :	
Dd	604	EATPSVLPDVI MREAPILNASVPSAGASVTPQSSPSLEASSVNYESI KHEPENPPPYEAM	663

Query Match 2.3%; Score 178.5; DB 10; Length 1192;
Best Local Similarity 19.4%; Pred. NO. 0.0067;
Matches 165; Conservative 114; Mismatches 302; Indels 269; Gaps 36;

QY	34	QOOQPPPPPPPPPPPPPPQAQLLPQPQQPPPPPPPPPPPPGPAVABEPLRRPKKELSA	93
Db	135	EDDEPPARPPPPPPASV---SPQAEPVMTTPAPAPAAPPSTPA-----APKRGGSS	182
QY	94	TKDVRNHCILTCENITVAOSVRNSPEFOKLGLIAMELFLICDDDAESDVRMVAOECNLKV	153
Db	183	GSVDETLFPALPRASEPVIIRSSAENDLKEQONTTISA-----QOEDFPSVLLETA---	232
QY	154	IKALMDSNPRLQOLELYKEIKKNAGPRSLRAALWRFELAHLVRPKCRKYLVNLLPCLT	213
Db	233	-----ASLPUSL-----PLSAAASFKEH-----YLGNLTSTVLP	260
QY	214	RTSKRPEESVQETLAAAQVPKIMAFSGNFANDNEIKVLKFIANTLANKSS---SPATRRATAA	270
Db	261	-----TEGTQLENVSFASKEVSEKATLLIDRLDTFSELEYSEMGSFSFVSP-----KA	310
QY	271	GSAVSCIQHSRTOFYFSYLLNVLLGLLPVPEDEHSTLLIIGVLLTYRYLWPLLOQOVKD	330
Db	311	ESAVIVA-----NPREIVKNKDEEEKLVSNMLHNQOGEPLTALTUKLVE	356
QY	331	TSL-----KGSGFGVTRKMEVSPSAEQVLQVYEILTTHHTQHODHVNYVTGALELLOOLFR	384


```

Db 357 DEVVSEKAKDSNEKRVAVE-APMREYADF----- 387
Qy 385 TTPPELLQTLTAVGGIGOLTAKEESGGRSRSIVELIAGGSSCPVLSRKQKGVLL 444
Db 388 -KPE-----RYVEKDKSDS-----DMLAAGGKTESNLESKVDK-KCFA 426
Qy 445 G--EEEALEDDESRSDVSSALT-ASVKDEISGEL-----AASSGVST-----P 486
Db 427 DSLEQINHEKDSSESSNDTSFSTPGIKDRPCAYITCAPFNPAAATESIATNIFPLGDP 486
Qy 487 GSAGH-----DIITEQPRSOHT-----LOADSVDLASCOTSSA 520
Db 487 TSENKTDKIEKKAQIVTEKNTSTKSNPFLVAAQDSETDYVTITDNLTKVTEEVVAMM 546
Qy 521 TGDDEEDILSHS--SSOVSAPVS-----DPAMDNDGTOASSPLSDS-----SOTTTEGPD 570
Db 547 PEGLTDLVQEAESSELENEVTGKIAYETKMDL---VQTEVWQESLYPAAQLCPSFEES 603
Qy 571 AVTPSD--SSSIVLDGTDNOYL---GLQIGQPODEDEATGI----- 607
Db 604 EATPSPVLPDVMEAPLNSAVPSAGASVIOPSSSPLEASSVNYESIKHEPNPPPYEEAM 663
Qy 608 -----LPDASAFNRSSMALQAH-----LLKNMHCRCQPSDSVVKFVL- 648
Db 664 SVSLKKVSGIKKEIKEP-ENINAAQETEPAYISIACDLIKETKLSAEPAPDFSDYSEMA 722
Qy 649 -----RDEATEPGDQENKPCRKINGDQSDDDASAPLV-----HCVRLLLSASFLLTG 695
Db 723 KVEQVPDHSSELVEDSPSEVDL-----FSDSDIIPVQKQDETVMVKESLITERS 775
Qy 696 GKNVL-----VPDRDVRVSVKALASCVGAVALHPE--SPFSKLYKVPPLDTTEY 743
Db 776 FESMIEYENKEKLSALPPEGKPYLESFKLSLDNTKDTLLPDBVSPLSKKEIPLQMEEL 835
Qy 744 PEOQVSDIL 753
Db 836 STAVYSNDDL 845

```

RESULT 9

```

US-10-108-605-29
; Sequence 29, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-29

```

```

Query Match 2.2%; Score 177; DB 9; Length 1404;
Best Local Similarity 19.2%; Pred. No. 0.01;
Matches 170; Conservative 118; Mismatches 327; Indels 272; Gaps 37;

```

```

Qy 83 PLH-RPKKELSATKK-----DRVNHCLTICENIV-AQSVRNSPEFQKLLGTAME 129
Db 307 PLHTRPTIYSVTAKVRGNPKALIIDIIVTQAGTYIKELVHGEGRTTSPSLSSICKPMD 366

```

```

Qy 130 LFLCSDDAESDVRWVADCLNKV--IKALMDSNLPRLQLELYKEIKKN-----GAPRS 181
Db 367 IQALDVSITTESPAVHYISEDSESVTYVSGAVDCEV-----KDSNNQNSAVGAGAPLP 418
Qy 182 LRAALWFAEL-----AHLVRPKCRPYLVNLLPCLTTRTSKR---PEESVOETL 227
Db 419 AIMKIORAEAPIGAVQLTPLWEHILTRRL-PETIFPSAMYAEFHERLQDPQEWQROHA 477
Qy 228 AAAPVKIMASFGNANDN-EIKVLKAFIANLKSSSPTRRTA-----AGSAVSIQ 278
Db 478 LRVLDVLVWQDDADGMEREQILGLVEMLHQATVRKGAIDCLRLVLAETAP- 534
Qy 279 HSRRTQFYFWLLNVLLGLL-----VPVEDEHSTLLILGVLLTLRLYVLPLOQOQVKT 331
Db 535 -----ETVMLGILDAGLTQKVPADSEHMGRLSCGVLLS-----VPALLQSIHMT 578
Qy 332 SLKSGFGVTRKEMEVSFAEQLVQVYELTLHHTQHODHNVTGALELLQQLFRTPPELL 391
Db 579 A-----QRHRIVHSTVERV-----AHMDQVVOQEIIVKVLISKIRELLGVHEFEI 624
Qy 392 -----OTLTAVGGIGOL--TAAKEESGSRSGSIVELIAGG----- 427
Db 625 MGNVGRGDTLSRYQLSOVYGVSGKSKTGGEVKAGARALPREQGWGRNGNAVQNLDITLQ 684
Qy 428 --SSCSPVLSRKQKGVLLGEEALEDD-----SESRDSVSSALTAS 468
Db 685 IOSNCLDM-----GKIIMETEIKINDDTVTMRILEADTETEDGTHQEPNIGAGIVKVI 738
Qy 469 VKDEISGELAASSGVSTPGSA-----CHDII-----TEQPRS-----QH 502
Db 739 SDSELDPEHIIKAGSVSEPTPSRGLKRVTFGEIVKMTPSDASSTNNSRSTQTNQS 798
Qy 503 TLQADSVDLASCDLTSSATDGEDILS-----HSSQSVSAVPSDPAIDL-NDGTOASS 555
Db 799 ITVSSSEDAALSDLVSS-----DEKSTLSRPVNDKRTALVLEIFDNTKPLPQDRSLCSQ 854
Qy 556 FIDSSTOTTTEGSAVTPSDSSEIVLDGTONQXVLGLOIGQPODEDEATGILPDEASEA 615
Db 855 PHRHQSNQASE-----PSSNS-----PRNRQENKARLSPSPSPG 889
Qy 616 PRNSS-----NALQQAHLKKNMHCROPSSDSVDKFLVRDEATEPGQENKPC 663
Db 890 FKRSTSTPGTGNNTSPKVPKHQIEVLHNLQ--RDPSPRS-----QRSEDMGNDGKAL 942
Qy 664 RIKGDIGQSTDDDSAPLVHVCVRLLSASFLLTGGKNVLVPRDRVRSVKA----- 712
Db 943 HATGNPSSS---PTQPLISRTSRASASTMSPVSPATPKSWEDLDI-VNLKTLLELRSGOWRN 998
Qy 713 -----LALSCVGAVALHP-----ESFFSKLYKVPPLDT 740
Db 999 RUMGIGOLELALSSSSNALVQPYLDSLRLTLSSERHFEVSDKLRELLVNLISRLPLDN 1058
Qy 741 TEYPEQVVSIDILNVIDHGDPOVGA--TAILCGTLICISILSRSRFH 785
Db 1059 LEERTPQILSGLCROGNAGANRVCKALMQRLPAGTIVAKLITSPEELH 1105

```

RESULT 10

```

US-10-062-254-240
; Sequence 240, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwon
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jjian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping

```



```

RESULT 14
US-09-946-805-4
; Sequence 4, Application US/09946805
; Patent No. US20020116734A1
; GENERAL INFORMATION:
; APPLICANT: Dickman, Martin B.
; TITLE OF INVENTION: PLANT DERIVED BAG HOMOLOGUES
; FILE REFERENCE: 480140.469
; CURRENT APPLICATION NUMBER: US/09/946,805
; CURRENT FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-946-805-4

Query Match          2.1%; Score 164; DB 10; Length 1043;
Best Local Similarity 20.3%; Pred. NO. 0.042;

```


[illegible]

Search completed: January 28, 2003, 14:40:27
Job time : 33.2305 secs

Db 541 AVPSDPMALNDGTQASSPISDSSQTTEGPDASVTPSDSSSEIVLDGTNDQYIGLQIGQP 600
Qy 598 QDEDEATGILPDEASEAFRNSMALQOAHLLKNSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSMALQOAHLLKNSHCRQPSDSSVDKFLVLRDEATEPGD 660
Qy 658 QENKPCRIGDGTGSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 661 QENKPCRIGDGTGSDDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 720
Qy 718 VGAVALHPESFFSKLYKVPLOTTEPEQYVSDILNYIDHGDQVVRGATATLCTGLICS 777
Db 721 VGAVALHPESFFSKLYKVPLOTTEPEQYVSDILNYIDHGDQVVRGATATLCTGLICS 780
Qy 778 ILSRFRHVGDMGTTTITLTGNTFSLADICPILRLTKLDESSVTCCLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGDMGTTTITLTGNTFSLADICPILRLTKLDESSVTCCLACTAVRNCVMSLC 840
Qy 838 SSSYSELGLQLIIDVLTTLNSSVWLVRTELTETLAEIDFLVSLFLEAKAENLHRCAGHYT 897
Db 841 SSSYSELGLQLIIDVLTTLNSSVWLVRTELTETLAEIDFLVSLFLEAKAENLHRCAGHYT 900
Qy 898 GLLKQERVLNNVVIHLLGDEDPVRHVAASLIRLVPLKFKYKCDQGGQADPVVAVARQDS 957
Db 901 GLLKQERVLNNVVIHLLGDEDPVRHVAASLIRLVPLKFKYKCDQGGQADPVVAVARQDS 960
Qy 958 SVYKLMLHETQPSHFSYSTIRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
Db 961 SVYKLMLHETQPSHFSYSTIRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
Qy 1018 RALTFCCEALCLSTAFVPCVWSLGHGCVPPPLSASDESRSKSCVGMATMILTLLSSAW 1077
Db 1021 RALTFCCEALCLSTAFVPCVWSLGHGCVPPPLSASDESRSKSCVGMATMILTLLSSAW 1080
Qy 1078 FPLDLAHOALILAGNLIAASAPKSLRWSWASEEAPNPAATKQEEVWPALGDRLVPMV 1137
Db 1081 FPLDLAHOALILAGNLIAASAPKSLRWSWASEEAPNPAATKQEEVWPALGDRLVPMV 1140
Qy 1138 EQLFSLHLLKVINICAHVLDVAPGPAIKALPSLTWPPSLPIRRKKGKEPEQASVPL 1197
Db 1141 EQLFSLHLLKVINICAHVLDVAPGPAIKALPSLTWPPSLPIRRKKGKEPEQASVPL 1200
Qy 1198 SPKKGSEASARSQDSGTSGVTTTSKSSLSGFSYHPLSYLKLHDVLKATHANKVTLIDLQN 1257
Db 1201 SPKKGSEASARSQDSGTSGVTTTSKSSLSGFSYHPLSYLKLHDVLKATHANKVTLIDLQN 1260
Qy 1258 STKFGFLRSALDVLISOILELATLQDICKVEEILGYLKSCFSREPMMATVCVQQLLKT 1317
Db 1261 STKFGFLRSALDVLISOILELATLQDICKVEEILGYLKSCFSREPMMATVCVQQLLKT 1320
Qy 1318 LFGTNLASQFDGLSSNPKSQGQRAQLGSSSVRPGLYHYCFMAYPTHETOALADASLRNM 1377
Db 1321 LFGTNLASQFDGLSSNPKSQGQRAQLGSSSVRPGLYHYCFMAYPTHETOALADASLRNM 1380
Qy 1378 VQAEQENDTSGWFDVLQKYSTOLKTLNLTSTVTKNRADKNAIHNRILFELPVIKALKQYTT 1437
Db 1381 VQAEQENDTSGWFDVLQKYSTOLKTLNLTSTVTKNRADKNAIHNRILFELPVIKALKQYTT 1440
Qy 1438 TTCVQKQKQVLDLALQVLQVRNVCLLDSQVTFIGFVLQKQFIEYVQGFRESEAIIPNIF 1497
Db 1441 TTCVQKQKQVLDLALQVLQVRNVCLLDSQVTFIGFVLQKQFIEYVQGFRESEAIIPNIF 1500
Qy 1498 FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1501 FFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1536

RESULT 6

US-08-556-419A-1

; Sequence 1, Application us/08556419A

; GENERAL INFORMATION:

; APPLICANT: Ross, Christopher A

; APPLICANT: Li, Xiao-Jiang

; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan H
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul F
; APPLICANT: Snyder, Solomon H
; TITLE OF INVENTION: Huntingtin-Associated Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD.
; STREET: 1001 G Street, NW suite 1100
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,419A
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 1107.52271
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-556-419A-1

Query Match 99.18; Score 7821.5; DB 9; Length 3144;

Best Local Similarity 99.88; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MATLEKLMKAFESLKSF---OQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57

Db 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 60

Qy 58 QPLLPPQPP 117

Db 61 QPLLPPQPP 120

Qy 118 PEFQKLLGIAMLEFLCSDDAESDVRWVADECLANKVIKALMDSNLPRLQLELYKEIKNG 177

Db 121 PEFQKLLGIAMLEFLCSDDAESDVRWVADECLANKVIKALMDSNLPRLQLELYKEIKNG 180

Qy 178 APRSLRAALWFAELAHVLRPQKCRPYLVNLLPCLITRTSKRPEESVQETLAAAVPKIMAS 237

Db 181 APRSLRAALWFAELAHVLRPQKCRPYLVNLLPCLITRTSKRPEESVQETLAAAVPKIMAS 240

Qy 238 FGNFANDNEIKVLLKAFATANKLSSSPTTRTAAGSAVSAVICOHSRRTQYFYSWLLNVLLGL 297

Db 241 FGNFANDNEIKVLLKAFATANKLSSSPTTRTAAGSAVSAVICOHSRRTQYFYSWLLNVLLGL 300

Qy 298 LVPVEDEHSTLLILGLVLTLYLVLPLOQQVKTSLKSGFVTRKMEVSPSAEQLVQVY 357

Db 301 LVPVEDEHSTLLILGLVLTLYLVLPLOQQVKTSLKSGFVTRKMEVSPSAEQLVQVY 360

Qy 358 ELTLHHTQHQHNVTGALELIQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRG 417

Db 361 ELTLHHTQHQHNVTGALELIQQLFRTPPPPELLQTLTAVGGIGQLTAAKESGGRSRG 420

Qy 418 SIVELIAGGGSCSPVLSRKQKGVLLGEEALEDDDSRSDSVSSSALTATSVKDEISEL 477

Db 421 STVELIAGGSCSPVLSRKQKGVLLGEEALEDDSESVDSSALTASVKDISEL 480
QY 478 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDIITEQPSQHTLQADSVDLASCDLTSSATGDDEEDILSHSSQVS 540
QY 538 AVPSOPAMDLNDGTQASSPISDSSOTTTEGPD SATVPDSSSEIVLDGTDNOYLGLOIGQP 597
Db 541 AVPSOPAMDLNDGTQASSPISDSSOTTTEGPD SATVPDSSSEIVLDGTDNOYLGLOIGQP 600
QY 598 QDEDEATGILPDEASEAFRNSMALQQAHLKNNSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKNNSHCRQPSDSSVDKFLVRDEATEPGD 660
QY 658 QENKPCRIGDITGQSTDDDSAPLVHCVRLLNSASFLTGGKNVLPDRVRVSVKALALSC 717
Db 661 QENKPCRIGDITGQSTDDDSAPLVHCVRLLNSASFLTGGKNVLPDRVRVSVKALALSC 720
QY 718 VGAVALHPESFESKLYKVPDLTTEYPEEQVYSDILNVIDHGDPOVGCATAILCGTLICS 777
Db 721 VGAVALHPESFESKLYKVPDLTTEYPEEQVYSDILNVIDHGDPOVGCATAILCGTLICS 780
QY 778 ILSRFRHVGMGTIRLTGNTFTSLADCIPLLRKTKDESSVTCCKLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGMGTIRLTGNTFTSLADCIPLLRKTKDESSVTCCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGQLIIDLVTLRNSSYWLVRTELLETAEIDFRLVSLFLEAKAENLHRAHHT 897
Db 841 SSSYSELGQLIIDLVTLRNSSYWLVRTELLETAEIDFRLVSLFLEAKAENLHRAHHT 900
QY 898 GLLKQLQERVNNVTHLLGDEDPVRVHVAASLRVLPKFKYKDOGOADPVAVARDQS 957
Db 901 GLLKQLQERVNNVTHLLGDEDPVRVHVAASLRVLPKFKYKDOGOADPVAVARDQS 960
QY 958 SVYLLMHETQPPSHFVSITIRYGYNLLPSTTDVTMENNLSRVIAAASHELITSTT 1017
Db 961 SVYLLMHETQPPSHFVSITIRYGYNLLPSTTDVTMENNLSRVIAAASHELITSTT 1020
QY 1018 RALTEGCEALCLLSTAFPVCVWSLGHGCVPPPLSASDESRSKCTVGMATMILLLSAW 1077
Db 1021 RALTEGCEALCLLSTAFPVCVWSLGHGCVPPPLSASDESRSKCTVGMATMILLLSAW 1080
QY 1078 PFLDLASODALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLVPMV 1137
Db 1081 PFLDLASODALILAGNLLAASAPKSLRSSWASBEENAPATKOEYVMPALGDRLVPMV 1140
QY 1138 EQLFSLHLKLVINICAHVLDVAPGPAIKAAPLSLTNPSPSLSPIRRKKEKEPGEQASVPL 1197
Db 1141 EQLFSLHLKLVINICAHVLDVAPGPAIKAAPLSLTNPSPSLSPIRRKKEKEPGEQASVPL 1200
QY 1198 SPKKGSEASASROSSTSGPVTTSKSSLSGSFYHLPSYKLDHVLKATHANYKVTLDLON 1257
Db 1201 SPKKGSEASASROSSTSGPVTTSKSSLSGSFYHLPSYKLDHVLKATHANYKVTLDLON 1260
QY 1258 STEKFGGLRSALDVLQILELATLQIDIGKCVBEILGYLKSCEFREPMMATVCVQQLLKT 1317
Db 1261 STEKFGGLRSALDVLQILELATLQIDIGKCVBEILGYLKSCEFREPMMATVCVQQLLKT 1320
QY 1318 LFGTNLASQFGLSNPSKSGRAQRGLGSSVRPGLHYCFMAYTHTFTQALADASLRNM 1377
Db 1321 LFGTNLASQFGLSNPSKSGRAQRGLGSSVRPGLHYCFMAYTHTFTQALADASLRNM 1380
QY 1378 VQAEQNTSGFDFVQKVSFTQKLTNTSVTKNRADKNAIHHNRLFEPLVVKALKQYTT 1437
Db 1381 VQAEQNTSGFDFVQKVSFTQKLTNTSVTKNRADKNAIHHNRLFEPLVVKALKQYTT 1440
QY 1438 TTCVOLQKQVLDLALQVLQVRVNYCLLDSQVFFIGFVKQFEYTEVGFQFRESEAIIPNIF 1497
Db 1441 TTCVOLQKQVLDLALQVLQVRVNYCLLDSQVFFIGFVKQFEYTEVGFQFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1533
|||||

Db 1501 FFLVLLSYERYHSKQIIGIPKIIQICDGMASGRKA 1536

RESULT 7

US-08-556-419B-1
Sequence 1, Application US/08556419B
GENERAL INFORMATION:
APPLICANT: Ross, Christopher A
APPLICANT: Li, Xiao-Jiang
APPLICANT: Li, Shi-Hua
APPLICANT: Sharp, Alan H
APPLICANT: Lanahan, Anthony
APPLICANT: Worley, Paul F
APPLICANT: Snyder, Solomon H
TITLE OF INVENTION: Huntingtin-Associated Protein
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, LTD.
STREET: 1001 G Street, NW suite 1100
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,419B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 1107.52271
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3144 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-556-419B-1

Query Match 99.1%; Score 7821.5; DB 9; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF--QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQ 60
QY 58 QPLLPPQPP 117
Db 61 QPLLPPQPP 120
QY 118 PEFQKLLGIAMELFLCSDDAESDVRMVADECLNKVIKALMDSNLRLOLELYKEIKNG 177
Db 121 PEFQKLLGIAMELFLCSDDAESDVRMVADECLNKVIKALMDSNLRLOLELYKEIKNG 180
QY 178 APRSLRAALWRAELAHVLRPQRCRPLYNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 181 APRSLRAALWRAELAHVLRPQRCRPLYNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240
QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSICQHSRRRTQYFYSWLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSICQHSRRRTQYFYSWLLNVLLGL 300

Qy 298 LVPVEDEHSTLLILGLVLTLLRYLVLPLQQVKDTSLKSGFVTRKMEVSPASOLVQVY 357
Db 301 LVPVEDEHSTLLILGLVLTLLRYLVLPLQQVKDTSLKSGFVTRKMEVSPASOLVQVY 360
Qy 358 ELTLHHTQHODHNVTGALQLLOQLFRTPPELLOTLTAVGGICOLTAAKEESGRSRG 417
Db 361 ELTLHHTQHODHNVTGALQLLOQLFRTPPELLOTLTAVGGICOLTAAKEESGRSRG 420
Qy 418 SIVELIAGGGSCSPVLSRKQKGVLLGEEAEDEDSERSDVSSSALTASVKDEISEL 477
Db 421 SIVELIAGGGSCSPVLSRKQKGVLLGEEAEDEDSERSDVSSSALTASVKDEISEL 480
Qy 478 AASSGVSTPGSAGHDITTEQPSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDITTEQPSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
Qy 538 AVPSDPMDLNQTQASSPIDSSQTTTEGPDASVTPSDSSEIVLDGTDNOYLGILQIGQP 597
Db 541 AVPSDPMDLNQTQASSPIDSSQTTTEGPDASVTPSDSSEIVLDGTDNOYLGILQIGQP 600
Qy 598 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSMALQQAHLKNNKSHCRQPSDSSVDKFLVRDEATEPGD 660
Qy 658 QENKPCRIGDQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 661 QENKPCRIGDQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 720
Qy 718 VGAVALHPESFESKLYKVPDTEPEQYVSDILNYIDHGDPOVGRATALLCGTLICS 777
Db 721 VGAVALHPESFESKLYKVPDTEPEQYVSDILNYIDHGDPOVGRATALLCGTLICS 780
Qy 778 ILRSRFRHVGDMWGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 781 ILRSRFRHVGDMWGTIRTLTGNTFSLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLC 840
Qy 838 SSSYSELGQLIIDVLTLLRNSYVLVRLTELETLAEIDFRLVFLAEAKENLHRCAGHYT 897
Db 841 SSSYSELGQLIIDVLTLLRNSYVLVRLTELETLAEIDFRLVFLAEAKENLHRCAGHYT 900
Qy 898 GLLKQERVLNNVTHLLGDEDPVRHVAASLIRLPKLFYKCDQGOADPVAVARQDS 957
Db 901 GLLKQERVLNNVTHLLGDEDPVRHVAASLIRLPKLFYKCDQGOADPVAVARQDS 960
Qy 958 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASVHELITSTT 1017
Db 961 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASVHELITSTT 1020
Qy 1018 RALTFCCEALCLLSTAFVPCVWSLGHGCVPLPSASDESRSKCTVGMATMILLSSAW 1077
Db 1021 RALTFCCEALCLLSTAFVPCVWSLGHGCVPLPSASDESRSKCTVGMATMILLSSAW 1080
Qy 1078 FPLDLSAHQDALLTAGNLAAAPKSLRSSWASEEAPAAKQEEVWPALGDRAALVPMY 1137
Db 1081 FPLDLSAHQDALLTAGNLAAAPKSLRSSWASEEAPAAKQEEVWPALGDRAALVPMY 1140
Qy 1138 EQLFSLHLKVINICAHVLDVDPAGPAKALPSTLNPSPSLPIRRKGEKEPEQASVPL 1197
Db 1141 EQLFSLHLKVINICAHVLDVDPAGPAKALPSTLNPSPSLPIRRKGEKEPEQASVPL 1200
Qy 1198 SPKKGSEASAAASQSDTSGPTVTSKSSSLGSGFYHLPSYLLKHDVLKATHANYKVTLIDLQ 1257
Db 1201 SPKKGSEASAAASQSDTSGPTVTSKSSSLGSGFYHLPSYLLKHDVLKATHANYKVTLIDLQ 1260
Qy 1258 STEKFGFLRSALDVLSQILELATLQDICKVVEILGYLKCSFREPMMATVCVQOQLLKT 1317
Db 1261 STEKFGFLRSALDVLSQILELATLQDICKVVEILGYLKCSFREPMMATVCVQOQLLKT 1320
Qy 1318 LFGTNLASQFDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1377
Db 1321 LFGTNLASQFDGLSSNPSKQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1380
Qy 1378 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNRILFPELVIKALKQYTT 1437

Db 1381 VQAEQENDTSGWFDVLQKYSTQLKNTLTSVTKNRADKNAIHNRILFPELVIKALKQYTT 1440
Qy 1438 TTCVQKQKQVLDLLAQLVQLRVNYCLLDSQVFTGFVVKQFIEVQGFRESAIIIPNIF 1497
Db 1441 TTCVQKQKQVLDLLAQLVQLRVNYCLLDSQVFTGFVVKQFIEVQGFRESAIIIPNIF 1500
Qy 1498 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1501 FFLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1536
RESULT 8
US-09-212-271-21
; Sequence 21, Application US/09212271
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lananhan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.77898
; CURRENT APPLICATION NUMBER: US/09/212, 271
; CURRENT FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 3144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-212-271-21
Query Match 99.1%; Score 7821.5; DB 16; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
Qy 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 60
Qy 58 QLLPQQPP 117
Db 61 QLLPQQPP 120
Qy 118 PEFQKLLGIAMELFLCDDAEDSVRVADCLNKVTKALMDSNLPRLQLELYKEIKNG 177
Db 121 PEFQKLLGIAMELFLCDDAEDSVRVADCLNKVTKALMDSNLPRLQLELYKEIKNG 180
Qy 178 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 181 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 240
Qy 238 FGNFANDNEIKVLLKAFITANKLKSSPIRRTAAGSAVSIQHSRRTQYFYSWLLNVLGL 297
Db 241 FGNFANDNEIKVLLKAFITANKLKSSPIRRTAAGSAVSIQHSRRTQYFYSWLLNVLGL 300
Qy 298 LVPVEDEHSTLLILGLVLTLLRYLVLPLQQVKDTSLKSGFVTRKMEVSPASOLVQVY 357
Db 301 LVPVEDEHSTLLILGLVLTLLRYLVLPLQQVKDTSLKSGFVTRKMEVSPASOLVQVY 360
Qy 358 ELTLHHTQHODHNVTGALQLLOQLFRTPPELLOTLTAVGGICOLTAAKEESGRSRG 417
Db 361 ELTLHHTQHODHNVTGALQLLOQLFRTPPELLOTLTAVGGICOLTAAKEESGRSRG 420
Qy 418 SIVELIAGGGSCSPVLSRKQKGVLLGEEAEDEDSERSDVSSSALTASVKDEISEL 477
Db 421 SIVELIAGGGSCSPVLSRKQKGVLLGEEAEDEDSERSDVSSSALTASVKDEISEL 480
Qy 478 AASSGVSTPGSAGHDITTEQPSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537

Db	481	AASSGVSTPGSAGNDIITEIQPRSOHTLQAOBSVDLIASCDLTSSATDGEDIEDILSHSSQVS	541
Qy	538	AVPSDPAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNOYLGLIQIGOP	597
Db	541	AVPSDPAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNOYLGLIQIGOP	600
Qy	598	ODEDEAATGILPDEASEAFRNSMALQAQHLLKNMHCRCQPSDSSVDKFFVLRDEATEPGD	657
Db	601	QDEDEAATGILPDEASEAFRNSMALQAQHLLKNMHCRCQPSDSSVDKFFVLRDEATEPGD	660
Qy	658	QENKPCRIRKIGIGOSTDDSDAPLVHCVRLLSASFLLTGGKNVLVDPDRVRSVKALALSC	717
Db	661	QENKPCRIRKIGIGOSTDDSDAPLVHCVRLLSASFLLTGGKNVLVDPDRVRSVKALALSC	720
Qy	718	VCAAVALHPESFFSKLYKVPDLTTEYPEEYQVSDILNYIDHGDPQVRGATAILCGTLICS	777
Db	721	VCAAVALHPESFFSKLYKVPDLTTEYPEEYQVSDILNYIDHGDPQVRGATAILCGTLICS	780
Qy	778	ILSRSRFHVGDMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC	837
Db	781	ILSRSRFHVGDMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC	840
Qy	838	SSSYSELGLQLIDVLTLRNSSYWLVRTELLETAEIDFRLVSFLEAKAENLHRGAHYT	897
Db	841	SSSYSELGLQLIDVLTLRNSSYWLVRTELLETAEIDFRLVSFLEAKAENLHRGAHYT	900
Qy	898	GLLKLOERVLNNVILHLLGDEDPVRHVAAASLIRLVPKLFYKCGOQOQADPVAVARQOS	957
Db	901	GLLKLOERVLNNVILHLLGDEDPVRHVAAASLIRLVPKLFYKCGOQOQADPVAVARQOS	960
Qy	958	SVYLLKLHMETQPPSHFSVSTIIRYGYNLLPSITDVTMNNLSRVAATAVSHELITSTT	1017
Db	961	SVYLLKLHMETQPPSHFSVSTIIRYGYNLLPSITDVTMNNLSRVAATAVSHELITSTT	1020
Qy	1018	RALTFGCEALCLLSTAPPVCINSLGWCHGVPPLSASDESRSKCTVGMATMILTLLSSAW	1077
Db	1021	RALTFGCEALCLLSTAPPVCINSLGWCHGVPPLSASDESRSKCTVGMATMILTLLSSAW	1080
Qy	1078	FPDLASAQODALILAGNLLAASAPKSLRSWSASEEENPAAATQBEVWPALGDRAVPMV	1137
Db	1081	FPDLASAQODALILAGNLLAASAPKSLRSWSASEEENPAAATQBEVWPALGDRAVPMV	1140
Qy	1138	EQLFSHLKLKVINICAHVLDVAPGPAIKAAPLSLTNPPLSLPIRRKGKEPEGEQASVPL	1197
Db	1141	EQLFSHLKLKVINICAHVLDVAPGPAIKAAPLSLTNPPLSLPIRRKGKEPEGEQASVPL	1200
Qy	1198	SPKKGSEASASQSDTSGPVTTSKSSSILGSPYHLPYSYLKLHDVLKATHANYKVTLDLQN	1257
Db	1201	SPKKGSEASASQSDTSGPVTTSKSSSILGSPYHLPYSYLKLHDVLKATHANYKVTLDLQN	1260
Qy	1258	STKFGFGLRSALDVLSQLLELATLQDICKCVVEEILGYLKCFSREPMWATVCVOQLLKT	1317
Db	1261	STKFGFGLRSALDVLSQLLELATLQDICKCVVEEILGYLKCFSREPMWATVCVOQLLKT	1320
Qy	1318	LFGTNLASQFDGLSSNPSPKSQRAQRLGSSSVRPGLYHYFCFMAPYTHFTQALADASLRNM	1377
Db	1321	LFGTNLASQFDGLSSNPSPKSQRAQRLGSSSVRPGLYHYFCFMAPYTHFTQALADASLRNM	1380
Qy	1378	VQAEQENDTSGWFDVLQKYSTOLKTNLTSVTKNRADKNAINHIRLFEPLVIKALKQYTT	1437
Db	1381	VQAEQENDTSGWFDVLQKYSTOLKTNLTSVTKNRADKNAINHIRLFEPLVIKALKQYTT	1440
Qy	1438	TTCVQLQKQVLDLALVOLRVNYNCLLSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIF	1497
Db	1441	TTCVQLQKQVLDLALVOLRVNYNCLLSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIF	1500
Qy	1498	FFVLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA	1533
Db	1501	FFVLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKA	1536

RESULT 9

US-09-538-092-1118

```

: Sequence 1118, Application US/09538092
: GENERAL INFORMATION:
: APPLICANT: Giot, Loic
: APPLICANT: Mansfield, Traci A.
: TITLE OF INVENTION: Protein-Protein Complexes and Methods
: FILE REFERENCE: 15966-542
: CURRENT APPLICATION NUMBER: US/09/538,092
: CURRENT FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: 60/121,352
: PRIOR FILING DATE: 1999-04-01
: PRIOR APPLICATION NUMBER: 60/178,965
: PRIOR FILING DATE: 2000-02-01
: NUMBER OF SEQ ID NOS: 1387
: SOFTWARE: CurapatSeqformatter Version 0.9
: SEQ ID NO 1118
: LENGTH: 3144
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (0)...(0)
: OTHER INFORMATION: Polypeptide Accession Number P42858
: US-09-538-092-1118

```

Query Match	99.1%	Score 7821.5	DB 19	Length 3144
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1533	Conservative	Mismatches 0	Indels 3	Gaps 1

RESULT 9

US-09-538-092-1118

QY 1138 EQFSLHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGKEKEGEQASVPL 1197
DB 1220 EQFSLHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGKEKEGEQASVPL 1279
QY 1198 SPKKGSEASASQSDTSGPVTTSKSSLSGSFYHLPSYKLKLDVLAHANYKVTLDLQ 1257
DB 1280 SPKKGSEASASQSDTSGPVTTSKSSLSGSFYHLPSYKLKLDVLAHANYKVTLDLQ 1339
QY 1258 STEKFGFLRSALDVLISQILELATLQDICKVEEILGILKSCFSPREPMATVVCVQQLKT 1317
DB 1340 STEKFGFLRSALDVLISQILELATLQDICKVEEILGILKSCFSPREPMATVVCVQQLKT 1399
QY 1318 LFGTNLASQFDGLSSNPSPKSGQRAQLGSSVPRGLYHYCFMAYTHFTQALADASLRNM 1377
DB 1400 LFGTNLASQFDGLSSNPSPKSGQRAQLGSSVPRGLYHYCFMAYTHFTQALADASLRNM 1459
QY 1378 VQAEQENDTSGWFDVLQKSTQTKNLTSVTKNRADKNAIHNIHRIEPLVIAKALQYTT 1437
DB 1460 VQAEQENDTSGWFDVLQKSTQTKNLTSVTKNRADKNAIHNIHRIEPLVIAKALQYTT 1519
QY 1438 TTCVOLQKQVLDLALQVLQVRVNYCLLDSQVFIGFVLKQFIEVQGFRESAIIPINIF 1497
DB 1520 TTCVOLQKQVLDLALQVLQVRVNYCLLDSQVFIGFVLKQFIEVQGFRESAIIPINIF 1579
QY 1498 FFVLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
DB 1580 FFVLVLLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1615

RESULT 12

US-08-085-000-6

; Sequence 6, Application US/08085000

; GENERAL INFORMATION:

; APPLICANT: Macdonald, Marcy E.

; APPLICANT: Ambrose, Christine M.

; APPLICANT: Duyao, Mabel P.

; APPLICANT: Gusella, James F.

; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1225 Connecticut Avenue

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/085,000

; FILING DATE: 19930701

; CLASSIFICATION: 514

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 466-0800

; TELEFAX: (202) 833-8716

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3144 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-085-000-6

Query Match

Best Local Similarity 98.9%; Score 7805.5; DB 4; Length 3144;

Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY

1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57

DB 1 MATLEKLMKAFESLKSFOQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 60
QY 58 OPLLPQOPPPQA 117
DB 61 OPLLPQOPPPQA 120
QY 118 PEFQKLLGIAMELFLCSDDAESDVRVMADECLNKVKIKALMDSNLPRLQLELYEIKKNG 177
DB 121 PEFQKLLGIAMELFLCSDDAESDVRVMADECLNKVKIKALMDSNLPRLQLELYEIKKNG 180
QY 178 APRSRALWRAFAELAHVLPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAPVKIMAS 237
DB 181 APRSRALWRAFAELAHVLPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAPVKIMAS 240
QY 238 FGNFANDNEIKVLLKAFIANKLSSSPTIRRTAAGSAVSICQHSRRTQVFSWLLNVLGL 297
DB 241 FGNFANDNEIKVLLKAFIANKLSSSPTIRRTAAGSAVSICQHSRRTQVFSWLLNVLGL 300
QY 298 LVPVEDEHSTLLILGVLLTLRYLVPVLLQOQVKDTSLKSGFVTRKMEVSPSASOLVQVY 357
DB 301 LVPVEDEHSTLLILGVLLTLRYLVPVLLQOQVKDTSLKSGFVTRKMEVSPSASOLVQVY 360
QY 358 ELTLHTHQHNDHNVYTGALQLLQQLFRTPPELLOTLTAVGGIGOLTAAKESGGRSSG 417
DB 361 ELTLHTHQHNDHNVYTGALQLLQQLFRTPPELLOTLTAVGGIGOLTAAKESGGRSSG 420
QY 418 SIVELIAGGGSSCPVLSRKQKQKVLGEEALEDDSESRSVSSSALTASVKDEISGEL 477
DB 421 SIVELIAGGGSSCPVLSRKQKQKVLGEEALEDDSESRSVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCOLTSSATDGDDEEDILSHSSQVS 537
DB 481 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCOLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDPAMDNDGTQASSPISDSSTTTTEGPD SATVPDSSSEIVLDGTDNOYVLGIQIQP 597
DB 541 AVPSDPAMDNDGTQASSPISDSSTTTTEGPD SATVPDSSSEIVLDGTDNOYVLGIQIQP 600
QY 598 QDEDEATGILPDEASEAFRNSMALQQAHLKLNKSHCRQPSDSSVDKFLVLRDEATEPGD 657
DB 601 QDEDEATGILPDEASEAFRNSMALQQAHLKLNKSHCRQPSDSSVDKFLVLRDEATEPGD 660
QY 658 QENKPCRIKGDIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDROVRSVKALALSC 717
DB 661 QENKPCRIKGDIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDROVRSVKALALSC 720
QY 718 VGAVALHPESFESKLYKVPDTEYPEQYVSDILNYIDHGDPOVRGATAILCCTLICS 777
DB 721 VGAVALHPESFESKLYKVPDTEYPEQYVSDILNYIDHGDPOVRGATAILCCTLICS 780
QY 778 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
DB 781 ILSRSRHFVGDWMTGIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQIIDLVTLRNSSYVWVLTETLETFLRDLVSELEAKAENLRHAHHT 897
DB 841 SSSYSELGLQIIDLVTLRNSSYVWVLTETLETFLRDLVSELEAKAENLRHAHHT 900
QY 898 GLLKQERVNLNNVVIHLLGDEDPVRVHVAASLIRLVPKLFYKCDQGOQADPVAVARDQS 957
DB 901 GLLKQERVNLNNVVIHLLGDEDPVRVHVAASLIRLVPKLFYKCDQGOQADPVAVARDQS 960
QY 958 SVYLKLLMHETQPPSHFSVSTIRYRGNLPLPSITDVTMNNLSRVIAAVSHELITSTT 1017
DB 961 SVYLKLLMHETQPPSHFSVSTIRYRGNLPLPSITDVTMNNLSRVIAAVSHELITSTT 1020
QY 1018 RALTFCCEALCLLSTAFVPCVWLSGWHGCVPLPSASDESRSCTVGMATMILTLLSSAW 1077
DB 1021 RALTFCCEALCLLSTAFVPCVWLSGWHGCVPLPSASDESRSCTVGMATMILTLLSSAW 1080
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEBANPAATKQEEVWPALGDRALVPWV 1137
DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEBANPAATKQEEVWPALGDRALVPWV 1140

[illegible]

RESULT 14

```

14
US-08/724-152-8
; Sequence 8, Application US/08724152
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael
; APPLICANT: Lin, Biayang
; APPLICANT: Nasir, Jamal
; TITLE OF INVENTION: MOUSE MODEL FOR HUNTINGTON'S
; TITLE OF INVENTION: DISEASE AND
; TITLE OF INVENTION: RELATED DNA SEQUENCES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Bell, Seltzer, Park & Gibson
; STREET: 1211 East Morehead Street,
; CITY: Charlotte,
; STATE: North Carolina
; COUNTRY: U.S.A.
; ZIP: 34009
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,152
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/127,971
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Layton, Jr., Samuel G
; REGISTRATION NUMBER: 22807
; REFERENCE/DOCKET NUMBER: 3477-85
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 704-331-6000
; TELEFAX: 704-334-2114

```


QY 780 SRSRHFVGDWGTIRLTGNTESLADCIPLRLKTLKDESSVTCKLACTAVRNCVMSLSS 839
Db 760 SRSRLRVGDMGLNIRLTGNTESLVDICPLQLKTLKDESSVTCKLACTAVRNCVMSLSS 819
QY 840 SYSEGLQLIIDLVLNRNSYWLVRTELETLAEIDFRVLSPLEAKAENLHRGAHHYTG 899
Db 820 SYSDGLQLIDMLPLKNSYWLVRTELETLAEIDFRVLSPLEAKAENLHRGAHHYTG 879
QY 900 LKQERVLNNVYIHLGDEDPVRHVAASLRLVPKLFYKCDQAGQADPVAVARDQSSV 959
Db 880 LKQERVLNNVYIHLGDEDPVRHVAATSLRLVPKLFYKCDQAGQADPVAVARDQSSV 939
QY 960 YKLMLHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 1019
Db 940 YKLMLHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTRA 999
QY 1020 LTFGCEALCLLSTAFFVCINSLGHGCVPPPLSASDESKRKTGYGMATMILTSSAWFP 1079
Db 1000 LTFGCEALCLLSAAPPVCTNSLGHGCVPPPLSASDESKRKTGYGMATMILTSSAWFP 1059
QY 1080 LDLSAHQDALILAGNLLAASAPKLSRSSWASSEANPAATKOEYVWPALGDRALVPMVEQ 1139
Db 1060 LDLSAHQDALILAGNLLAASAPKLSRSSWSEANPAATKOEYVWPALGDRALVPMVEQ 1119
QY 1140 LFSHLLKVINICAHVLDVAPGPAIKAAALPSLTNPSPISPIRRKGKEKEGQASVPLSP 1199
Db 1120 LFSHLLKVINICAHVLDVTPGPAIKAAALPSLTNPSPISPIRRKGKEKEGQASVPLSP 1179
QY 1200 KGSASASASROSDTSGPVTSKSSLSGFSYHPLSYLKLHDVYLKATHANYKVTLDQNST 1259
Db 1180 KVGASASASROSDTSGPVTSKSSLSGFSYHPLSYLKLHDVYLKATHANYKVTLDQNST 1239
QY 1260 EKFGFSLRSALDVLSQLLELATLQIDGKVEEILGKSCFSPREPMATVCVQQLKTLF 1319
Db 1240 EKFGFSLRSALDVLSQLLELATLQIDGKVEEILGKSCFSPREPMATVCVQQLKTLF 1299
QY 1320 GTNLASQFDGLSSNPKSQGRAQLRGSSSVRPLGYHYCFMAYPTHFTQALADASILRNMQ 1379
Db 1300 GTNLASQFDGLSSNPKSQGRAQLRGSSSVRPLGYHYCFMAYPTHFTQALADASILRNMQ 1359
QY 1380 AEQENDTSQFVDLQKVSQTLKNTLTSVTKNRADKNAIHNIRLPEPLVIKALKQYTTT 1439
Db 1360 AEQENDTSQFVDLQKVSQTLKNTLTSVTKNRADKNAIHNIRLPEPLVIKALKQYTTT 1419
QY 1440 CVQLQKQVLDLQAQLVQVRYNCLDSQVFGFVLKQFIEVQGFRESEALIPNIEFF 1499
Db 1420 SVQLQKQVLDLQAQLVQVRYNCLDSQVFGFVLKQFIEVQGFRESEALIPNIEFF 1479
QY 1500 LVLLSYRHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1480 LVLLSYRHSKQIIGIPKIIQLCDGIMASGRKA 1513

RESULT 15

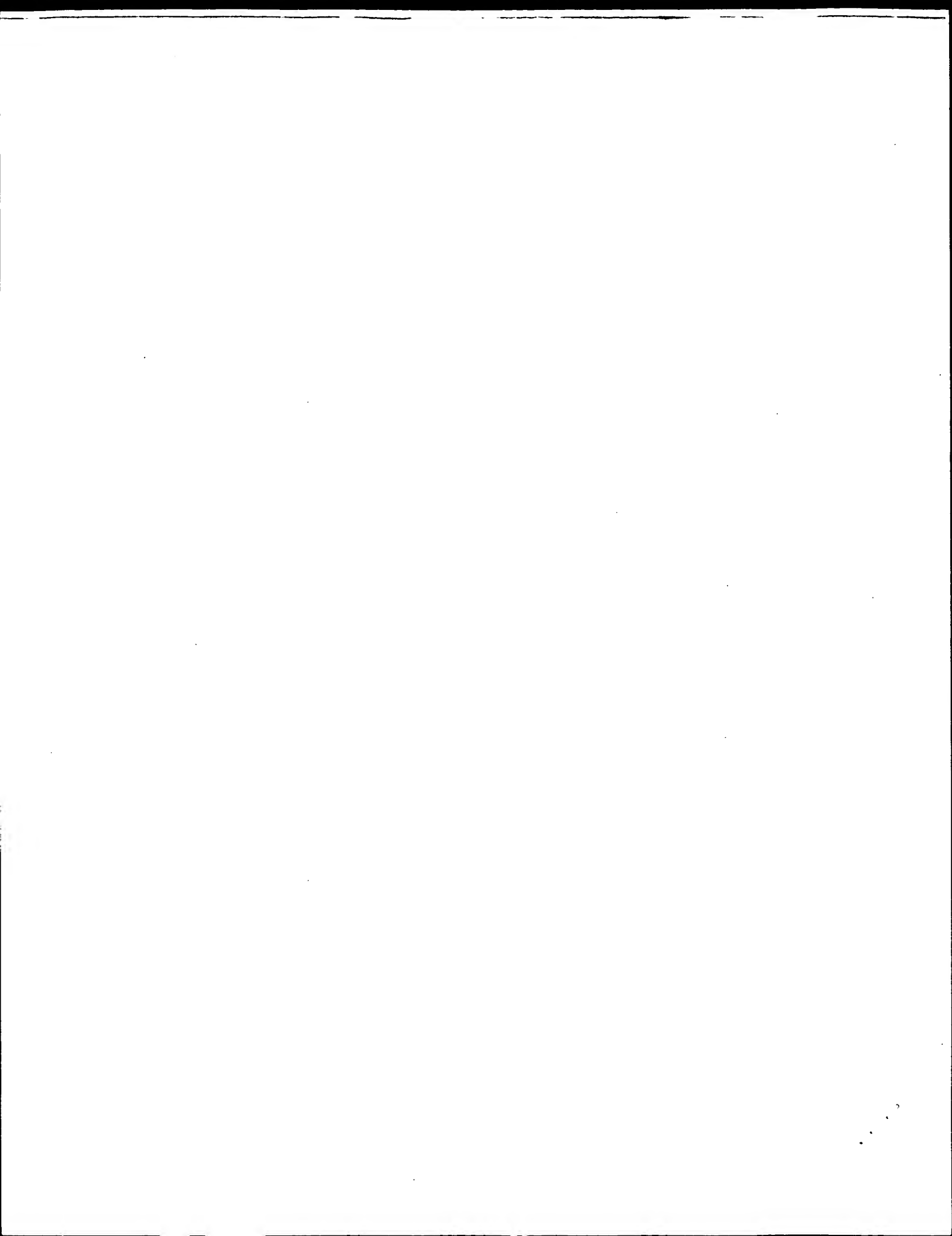
PCT-US99-05250-31
; Sequence 31, Application PC/TUS9905250
; GENERAL INFORMATION:
; APPLICANT: The Burnham Institute
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US99/05250
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/041,886
; FILING DATE: 12-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: FP-LJ 3484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US99-05250-31

Query Match 37.6%; Score 2967.5; DB 1; Length 589;
Best Local Similarity 99.5%; Pred. No. 7e-210;
Matches 586; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSP---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSPQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 60
QY 58 QPLLPPQPP 117
Db 61 QPLLPPQPP 120
QY 118 PEFQKLLGIAAMELFLCSDDAESDVRMVADECLNKVIALKALMDSNLPRLQLEYKEIKNG 177
Db 121 PEFQKLLGIAAMELFLCSDDAESDVRMVADECLNKVIALKALMDSNLPRLQLEYKEIKNG 180
QY 178 APRSLRAALRRFAELAHVLPQKCRPYLVNLLPCLTFTSKRPEESVQETLAAAVPKTMAS 237
Db 181 APRSLRAALRRFAELAHVLPQKCRPYLVNLLPCLTFTSKRPEESVQETLAAAVPKTMAS 240
QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAGSVAISICQHSRRTQYFYSWLLNVLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAGSVAISICQHSRRTQYFYSWLLNVLGL 300
QY 298 LVPVEDEHSTLLILGVLLTLRYLVPLQQQVKDTSLKGSFGVTRKEMEVSAPSABLQVY 357
Db 301 LVPVEDEHSTLLILGVLLTLRYLVPLQQQVKDTSLKGSFGVTRKEMEVSAPSABLQVY 360
QY 358 ELTLHHTQHODHNVTGAELELQQLFRTPPPELLOTLTAVGGIGQLTAAKEESGGRSRG 417
Db 361 ELTLHHTQHODHNVTGAELELQQLFRTPPPELLOTLTAVGGIGQLTAAKEESGGRSRG 420
QY 418 SIVELIAGGSSCSPVLSRKQKGVLLGEBEALDDSESRSDVSSSALTASVXDEISGEL 477
Db 421 SIVELIAGGSSCSPVLSRKQKGVLLGEBEALDDSESRSDVSSSALTASVXDEISGEL 480
QY 478 AASGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
Db 481 AASGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDPAIDLNDGTQASSPISDSSQTTTEGPDSPASVTPSDSEIYLDGTD 586
Db 541 AVPSDPAIDLNDGTQASSPISDSSQTTTEGPDSPASVTPSDSEIYLDGTD 589

Search completed: January 28, 2003, 14:37:35
Job time : 407.975 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on: January 28, 2003, 14:16.13 ; Search time 45.0972 Seconds
        (without alignments)
        2701.343 Million cell updates/sec
```

Title: US-09-904-987-7

Perfect score:

Sequence: 1 MATLEKLMKAFESLSKFQQQ.....DGIMASGRKASQPYRLCSP 1543

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 394613 seqs, 78952108 residues

Total number of hits satisfying chosen parameters: 394613

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*

```
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pcp.*
```

2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB pep: *

3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pcp:*

```
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pcp:*
5: /cgn2_6/ptodata/1/paa/US00_NEW_COMB.pcp:*
```

```
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB pep.*
```

```
6: /cyliz_6/ptodata/1/paa/us10_NEW_COMB.pcp: **
7: /cqn2_6/ptodata/1/paa/us60_NEW_COMB.pcp: **
```

1: /cguiz_v/pconata/1/paa/usuv_NEW_COMB.per:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	7821.5	99.1	2182	5	US-09-724-676-51187	Sequence 51187, A
2	7821.5	99.1	2182	5	US-09-724-676A-51187	Sequence 51187, A
3	7821.5	99.1	2324	5	US-09-724-676-51186	Sequence 51186, A
4	7821.5	99.1	2324	5	US-09-724-676A-51186	Sequence 51186, A
5	7821.5	99.1	3223	6	US-10-276-774-1777	Sequence 1777, Ap
6	7821.5	99.1	3223	6	US-10-276-774-1840	Sequence 1840, Ap
7	7086.5	89.8	1433	5	US-09-724-676-51188	Sequence 51188, A
8	7086.5	89.8	1433	5	US-09-724-676A-51188	Sequence 51188, A
9	6008	76.1	1763	5	US-09-724-676-51192	Sequence 51192, A
10	6008	76.1	1763	5	US-09-724-676A-51192	Sequence 51192, A
11	6008	76.1	1940	5	US-09-724-676-51191	Sequence 51191, A
12	6008	76.1	1940	5	US-09-724-676A-51191	Sequence 51191, A
13	5484	69.5	2690	5	US-09-724-676-51190	Sequence 51190, A
14	5484	69.5	2690	5	US-09-724-676A-51190	Sequence 51190, A
15	5273	66.8	1049	5	US-09-724-676-51181	Sequence 51181, A
16	5273	66.8	1049	5	US-09-724-676A-51181	Sequence 51181, A
17	2164.5	27.4	489	5	US-09-724-676-51189	Sequence 51189, A
18	2164.5	27.4	489	5	US-09-724-676A-51189	Sequence 51189, A
19	488	6.2	87	6	US-10-215-432-27	Sequence 27, Appl
20	373	4.7	68	6	US-10-215-432-28	Sequence 28, Appl
21	200.5	2.5	780	5	US-09-770-689B-5	Sequence 5, Appl
22	195.5	2.5	159	5	US-09-724-676-60613	Sequence 60613, A
23	195.5	2.5	159	5	US-09-724-676A-60613	Sequence 60613, A
24	195.5	2.5	2496	5	US-09-724-676-91031	Sequence 91031, A
25	195.5	2.5	2496	5	US-09-724-676A-91031	Sequence 91031, A
26	195.5	2.5	2548	5	US-09-724-676-91030	Sequence 91030, A

ALIGNMENTS

```

RESULT 1
US-09-724-676-51187
; Sequence 51187, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51187
; LENGTH: 2182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51187

```

Query Match 99.1%; Score 7821.5; DB 5; Length 2182;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

[illegible]


```
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEERALEDDSESRSDVSSSALTASVKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEERALEDDSESRSDVSSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCCLTSSATDGBEDILSHSSQVS 537
Db 523 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCCLTSSATDGBEDILSHSSQVS 582
QY 538 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQGP 597
Db 583 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQGP 642
QY 598 QDEDEATGILPDEASEAFRNSSMALQOAHLLKNMHCROQPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEATGILPDEASEAFRNSSMALQOAHLLKNMHCROQPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 703 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 762
QY 718 VGAVALHPESFESKLYKVPLOTTETPEQYVSDIILYIDHGDPOVGRGATATLTCGLTICS 777
Db 763 VGAVALHPESFESKLYKVPLOTTETPEQYVSDIILYIDHGDPOVGRGATATLTCGLTICS 822
QY 778 ILSRFRHVGDMGTIRTITGTNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 823 ILSRFRHVGDMGTIRTITGTNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGQLIIDVLTNRSSVWLVRTELLETLAEIDPRVLSFLEAKAENLHRAHHT 897
Db 883 SSSYSELGQLIIDVLTNRSSVWLVRTELLETLAEIDPRVLSFLEAKAENLHRAHHT 942
QY 898 GLKLOERVLNNVYIHLGDEDPVRHVAASLIIRLPKLFYKCDQGOADPVVAROQS 957
Db 943 GLKLOERVLNNVYIHLGDEDPVRHVAASLIIRLPKLFYKCDQGOADPVVAROQS 1002
QY 958 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASHELITSTT 1017
Db 1003 SVYKLMLHETOPPSHFSYSTIRYRGNLLPSITDVTMNNLSRVIAAASHELITSTT 1062
QY 1018 RALTFCCEALCLLSTAFVPCVWISLGHGCVPPPLSASDESRSCTVGMATMTLTLSSAW 1077
Db 1063 RALTFCCEALCLLSTAFVPCVWISLGHGCVPPPLSASDESRSCTVGMATMTLTLSSAW 1122
QY 1078 FPLDLSAHODALLTAGNLLAASAPKSLRSWASEEENPAATKQEVWPAIDGRALVPMV 1137
Db 1123 FPLDLSAHODALLTAGNLLAASAPKSLRSWASEEENPAATKQEVWPAIDGRALVPMV 1182
QY 1138 EQLFSHLLKVINICAHVLDVAPGPAIKAAPSLTNPPSLPIRRKGKEPEGEQASVPL 1197
Db 1183 EQLFSHLLKVINICAHVLDVAPGPAIKAAPSLTNPPSLPIRRKGKEPEGEQASVPL 1242
QY 1198 SPKKGSEASASQSDTSGPVTTSKSSSLGSPHPLSYLKLHDVLKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASASQSDTSGPVTTSKSSSLGSPHPLSYLKLHDVLKATHANYKVTLDLQN 1302
QY 1258 STEKFGFLRSALDVLSQLLELATLQDICKVEEILGLYKSCFSRPMMATVCVQQLLKT 1317
Db 1303 STEKFGFLRSALDVLSQLLELATLQDICKVEEILGLYKSCFSRPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPKSQGRLGSSSVRRPGLYHYCFMAYPTHFTQALADASLRNN 1377
Db 1363 LFGTNLASQFDGLSSNPKSQGRLGSSSVRRPGLYHYCFMAYPTHFTQALADASLRNN 1422
QY 1378 VQAEQENDTSGWFDVLQKYSTQLKTNLTSTVTKNRADKNAHNNHILFELPVTKALKQYTT 1437
Db 1423 VQAEQENDTSGWFDVLQKYSTQLKTNLTSTVTKNRADKNAHNNHILFELPVTKALKQYTT 1482
QY 1438 TTCVOLQKQVLDLQAQLVQLRVNYCCLDSQVFIGFVLKQFIEYVGOFRESEAIIPNIF 1497
Db 1483 TTCVOLQKQVLDLQAQLVQLRVNYCCLDSQVFIGFVLKQFIEYVGOFRESEAIIPNIF 1542
QY 1498 FFLVLSYERYHKSQIIGIKPIIQLCDGIMASGRKA 1533
```

```
Db 1543 FFLVLSYERYHKSQIIGIKPIIQLCDGIMASGRKA 1578
RESULT 2
US-09-724-676A-51187
; Sequence 51187, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51187
; LENGTH: 2182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51187
Query Match 99.1%; Score 7821.5; DB 5; Length 2182;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
Db 43 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 102
QY 58 QPLLPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PEFOKLGIAMELFLCDDAESDVRVVADECLNKVTKALMDSNLPRLQLELYKEIKNG 177
Db 163 PEFOKLGIAMELFLCDDAESDVRVVADECLNKVTKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSLRAALWRFABLAHLVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFIANKLKSSPTIRRTAAGSAVSICOHSSRTQYFYSWLLNLVLLGL 297
Db 283 FGNFANDNEIKVLLKAFIANKLKSSPTIRRTAAGSAVSICOHSSRTQYFYSWLLNLVLLGL 342
QY 298 LVPVEDEHSTLILGLVLTLLRYLPLVLLQQQVKDTSLKGSFCVTRKEMEVSFSAQLVQVY 357
Db 343 LVPVEDEHSTLILGLVLTLLRYLPLVLLQQQVKDTSLKGSFCVTRKEMEVSFSAQLVQVY 402
QY 358 ELTLHHTQHDHNVVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 417
Db 403 ELTLHHTQHDHNVVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 462
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEERALEDDSESRSDVSSSALTASVKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEERALEDDSESRSDVSSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCCLTSSATDGBEDILSHSSQVS 537
Db 523 AASSGVSTPGSAGHDIITEQPRSQHTLQADSVDLASCCLTSSATDGBEDILSHSSQVS 582
QY 538 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQGP 597
Db 583 AVPSDPAMDNDGTQASSPIDSSQTTTEGPDPSAVTPSDSSEIVLDGTDNOYLGIGIQGP 642
QY 598 QDEDEATGILPDEASEAFRNSSMALQOAHLLKNMHCROQPSDSSVDKFLVRDEATEPGD 657
Db 643 QDEDEATGILPDEASEAFRNSSMALQOAHLLKNMHCROQPSDSSVDKFLVRDEATEPGD 702
QY 658 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 717
Db 703 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSC 762
```


QY 718 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNIDHGDQVVRGATAILCGTLICS 777
Db 763 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNIDHGDQVVRGATAILCGTLICS 822
QY 778 ILSRSRHFVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVCKLACTAVRNCVMSLC 837
Db 823 ILSRSRHFVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSFLPAKAENLHRGAHHT 897
Db 883 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSFLPAKAENLHRGAHHT 942
QY 898 GLKLOERVNNVNIHLGDEDEPRVRHVAASLIRLVKLFYKCDQOGADPVVAVARDO 957
Db 943 GLKLOERVNNVNIHLGDEDEPRVRHVAASLIRLVKLFYKCDQOGADPVVAVARDO 1002
QY 958 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITSTT 1017
Db 1003 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITSTT 1062
QY 1018 RALTFCCEALCILLSTAFVPCINSLGWHGCVPLSLASDESRSKCTVGMATNLTLLSSAW 1077
Db 1063 RALTFCCEALCILLSTAFVPCINSLGWHGCVPLSLASDESRSKCTVGMATNLTLLSSAW 1122
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSWSASEEENPAATKQEBWVPALGDRLVPMV 1137
Db 1123 FPLDLSAHQDALILAGNLLAASAPKSLRSWSASEEENPAATKQEBWVPALGDRLVPMV 1182
QY 1138 EQLFSLHLKVINICAHVDDVAPGPAIKAAPSLTNPPSLPTIRRKGEKEPEGEQASVPL 1197
Db 1183 EQLFSLHLKVINICAHVDDVAPGPAIKAAPSLTNPPSLPTIRRKGEKEPEGEQASVPL 1242
QY 1198 SPKKGSPASASQSDTSGPVTTKSSSLGSGFYHPLSYLKLHDVLKATHANYKVTLDLQ 1257
Db 1243 SPKKGSPASASQSDTSGPVTTKSSSLGSGFYHPLSYLKLHDVLKATHANYKVTLDLQ 1302
QY 1258 STEKFGFLRSALDVLSQLLELATLQDQKVEEILGKLCSCFSRPPMMATVCVQQLLKT 1317
Db 1303 STEKFGFLRSALDVLSQLLELATLQDQKVEEILGKLCSCFSRPPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQPDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQPDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSGWFDVLQKYSTQLKTNLTSVTKNRADKNAIHNHIRLFEPLVTKALKQYTT 1437
Db 1423 VQAEQENDTSGWFDVLQKYSTQLKTNLTSVTKNRADKNAIHNHIRLFEPLVTKALKQYTT 1482
QY 1438 TTCVQLQKQVLDLLAQLVLRVNYCCLDSQVFIQVFLVQFIEYVQGFRESEAIIPNIF 1497
Db 1483 TTCVQLQKQVLDLLAQLVLRVNYCCLDSQVFIQVFLVQFIEYVQGFRESEAIIPNIF 1542
QY 1498 FFLVLLSYERYHSKQIGIPKIIQLCDGIMASGRKA 1533
Db 1543 FFLVLLSYERYHSKQIGIPKIIQLCDGIMASGRKA 1578

RESULT 3

US-09-724-676-51186
; Sequence 51186, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51186
; LENGTH: 2324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51186

Query Match 99.1%; Score 7821.5; DB 5; Length 2324;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MATLEKLMKAFESLSKF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKLMKAFESLSKFQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLQPOPP 117
Db 103 QPLLQPOPP 162
QY 118 PEFQKLLGIGAMELFLCSDDAESDVRVMADECLNKVIRKALMDSNLPRLQLELYKEIKNG 177
Db 163 PEFQKLLGIGAMELFLCSDDAESDVRVMADECLNKVIRKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSLRAALWRAFAELHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSLRAALWRAFAELHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 297
Db 283 FGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLGL 342
QY 298 LVPVEDEHSTLLILGVLTLRLVPLLOQVKOTSLKSGFVTRKEMEVSFAEOLVOY 357
Db 343 LVPVEDEHSTLLILGVLTLRLVPLLOQVKOTSLKSGFVTRKEMEVSFAEOLVOY 402
QY 358 ELTLHHTQHONHVVYTGALQLLQQLFRTPPPPELLQTLTAVGGIGQTLAAKESGGRSRG 417
Db 403 ELTLHHTQHONHVVYTGALQLLQQLFRTPPPPELLQTLTAVGGIGQTLAAKESGGRSRG 462
QY 418 SIVELIAGGGSCSPVLRRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISEL 477
Db 463 SIVELIAGGGSCSPVLRRKQKGVLLGEEALEDDSESDVSSSALTASVKDEISEL 522
QY 478 AASSGVTPGSAAGHDITTEQPSOHTLOADSVDLASCDLTSSATDGEDILSHSSSOVS 537
Db 523 AASSGVTPGSAAGHDITTEQPSOHTLOADSVDLASCDLTSSATDGEDILSHSSSOVS 582
QY 538 AVPSDAMDNDGTQASSPIDSSQTTTEGPD SATVTPSDSSSEIVLDGTDNQLGIIQGP 597
Db 583 AVPSDAMDNDGTQASSPIDSSQTTTEGPD SATVTPSDSSSEIVLDGTDNQLGIIQGP 642
QY 598 QDEDEEATGIUPDEASEAFNRSSMALQQAHLKNNHCHRCQPSDSSVDFKVLDEATEPGD 657
Db 643 QDEDEEATGIUPDEASEAFNRSSMALQQAHLKNNHCHRCQPSDSSVDFKVLDEATEPGD 702
QY 658 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGGKNVLPDROVRVSVKALALSC 717
Db 703 QENKPCRIGDQIGQSTDDSDAPLVHCVRLLSASFLLTGGKNVLPDROVRVSVKALALSC 762
QY 718 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNIDHGDQVVRGATAILCGTLICS 777
Db 763 VGAVALHPESFESKLYKVPDLTTEPEQYVSDILNIDHGDQVVRGATAILCGTLICS 822
QY 778 ILSRSRHFVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVCKLACTAVRNCVMSLC 837
Db 823 ILSRSRHFVGDWMTTIRTLTGNTFSLADCIPLLRKLTLDSESVCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSFLPAKAENLHRGAHHT 897
Db 883 SSSYSELGLQIIDLVTLRNSSVWLVRTLETLAEIDFRLVSFLPAKAENLHRGAHHT 942
QY 898 GLKLOERVNNVNIHLGDEDEPRVRHVAASLIRLVKLFYKCDQOGADPVVAVARDO 957
Db 943 GLKLOERVNNVNIHLGDEDEPRVRHVAASLIRLVKLFYKCDQOGADPVVAVARDO 1002
QY 958 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITSTT 1017
Db 1003 SVYLKLLMHETQPPSHFSVSTIRIYRGYNLLPSITDVTMNNLSRVIAAASHELITSTT 1062


```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51188

Query Match      89.8%; Score 7086.5; DB 5; Length 1433;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKMKAFESLSKFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PFQKLLGTAMELFLLCSDDAESDVRWVADECLNKVKALMDSNLPRLQLELYKEIKNG 177
Db 163 PFQKLLGTAMELFLLCSDDAESDVRWVADECLNKVKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSRAALWRAEALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSRAALWRAEALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFATANKSSSPTTRTAAGSAVSIQHSRRTQFYFWLLNVLLGL 297
Db 283 FGNFANDNEIKVLLKAFATANKSSSPTTRTAAGSAVSIQHSRRTQFYFWLLNVLLGL 342
QY 298 LVPVEDEHSTLILGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 357
Db 343 LVPVEDEHSTLILGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 402
QY 358 ELTLHHTQHQNHNVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 417
Db 403 ELTLHHTQHQNHNVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 462
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDESDRSVSSALTAASKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDESDRSVSSALTAASKDEISGEL 522
QY 478 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTASSATDGEEDILSHSSQVS 537
Db 523 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTASSATDGEEDILSHSSQVS 582
QY 538 AVSDPAMDNDGTQASPSIDSSQTTTEGPDASVTPSDSEIIVLDGTDNQYLGQLGQP 597
Db 583 AVSDPAMDNDGTQASPSIDSSQTTTEGPDASVTPSDSEIIVLDGTDNQYLGQLGQP 642
QY 598 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 643 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVLRDEATEPGD 702
QY 658 QENKPCRIGKDIGSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 717
Db 703 QENKPCRIGKDIGSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPQVRGATAILCGTLICS 777
Db 763 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPQVRGATAILCGTLICS 822
QY 778 ILSRSRPHVGDWGMTIRTLTGNTFSLADCIPLLRKTKLDESSVTCCKLACTAVRNCVMSLC 837
Db 823 ILSRSRPHVGDWGMTIRTLTGNTFSLADCIPLLRKTKLDESSVTCCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQLIIDVFLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLRGAAHYT 897
Db 883 SSSYSELGLQLIIDVFLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLRGAAHYT 942
QY 898 GLLKQSERVLNNVYIHLGDEDPVRHVAAASLIRLVKFLFYKCDQGDQADPPVAVARDQS 957
Db 943 GLLKQSERVLNNVYIHLGDEDPVRHVAAASLIRLVKFLFYKCDQGDQADPPVAVARDQS 1002
QY 958 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1017
Db 1003 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1062
QY 1018 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1077
Db 1063 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEENPAATKQEEYVWPALGDRALVPMV 1137
Db 1123 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEENPAATKQEEYVWPALGDRALVPMV 1182
QY 1138 EQLFSLHLLKVNINICAHVLDVAVCPAIPAALPSLTNPPSLSPIRRKKEKEPEGEQASVPL 1197
Db 1183 EQLFSLHLLKVNINICAHVLDVAVCPAIPAALPSLTNPPSLSPIRRKKEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASARSQSDTSGPVTTTSKSSSLGFSYHLPYSLKLDHVLKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASARSQSDTSGPVTTTSKSSSLGFSYHLPYSLKLDHVLKATHANYKVTLDLQN 1302
QY 1258 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1317
Db 1303 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPKSKSGRAORLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQFDGLSSNPKSKSGRAORLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388
Db 1423 VQAEQENDTSG 1433

RESULT 8
US-09-724-676A-51188
; Sequence 51188, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51188
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51188

Query Match      89.8%; Score 7086.5; DB 5; Length 1433;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 57
Db 43 MATLEKMKAFESLSKFQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPPPQA 102
QY 58 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 117
Db 103 QPLLPOQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 162
QY 118 PFQKLLGTAMELFLLCSDDAESDVRWVADECLNKVKALMDSNLPRLQLELYKEIKNG 177
Db 163 PFQKLLGTAMELFLLCSDDAESDVRWVADECLNKVKALMDSNLPRLQLELYKEIKNG 222
QY 178 APRSRAALWRAEALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 237
Db 223 APRSRAALWRAEALHVRPQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMAS 282
QY 238 FGNFANDNEIKVLLKAFATANKSSSPTTRTAAGSAVSIQHSRRTQFYFWLLNVLLGL 297
Db 283 FGNFANDNEIKVLLKAFATANKSSSPTTRTAAGSAVSIQHSRRTQFYFWLLNVLLGL 342
QY 298 LVPVEDEHSTLILGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 357
Db 343 LVPVEDEHSTLILGVLLTLRLVPLLOQQVKDTSLKSGFVTRKMEVSPSAEQLVQVY 402
QY 358 ELTLHHTQHQNHNVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 417
Db 403 ELTLHHTQHQNHNVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAKEESGGRSRG 462
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDESDRSVSSALTAASKDEISGEL 477
Db 463 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDESDRSVSSALTAASKDEISGEL 522
QY 478 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTASSATDGEEDILSHSSQVS 537
Db 523 AASGVSTPGSAGHDITEQPRSQHTLQADSVDLASCDLTASSATDGEEDILSHSSQVS 582
QY 538 AVSDPAMDNDGTQASPSIDSSQTTTEGPDASVTPSDSEIIVLDGTDNQYLGQLGQP 597
Db 583 AVSDPAMDNDGTQASPSIDSSQTTTEGPDASVTPSDSEIIVLDGTDNQYLGQLGQP 642
QY 598 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 643 QDEDEEATGILPDEASEAFNRSSMALQAHLLKNKSHCRQPSDSSVDKFLVLRDEATEPGD 702
QY 658 QENKPCRIGKDIGSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 717
Db 703 QENKPCRIGKDIGSTDDSDAPLVHCVRLLSASFLLTGKKNVLPDRDRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPQVRGATAILCGTLICS 777
Db 763 VGAVALHPESFFSKLYKVPDLTTEPEEQVSDILNYIDHGDPQVRGATAILCGTLICS 822
QY 778 ILSRSRPHVGDWGMTIRTLTGNTFSLADCIPLLRKTKLDESSVTCCKLACTAVRNCVMSLC 837
Db 823 ILSRSRPHVGDWGMTIRTLTGNTFSLADCIPLLRKTKLDESSVTCCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQLIIDVFLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLRGAAHYT 897
Db 883 SSSYSELGLQLIIDVFLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLRGAAHYT 942
QY 898 GLLKQSERVLNNVYIHLGDEDPVRHVAAASLIRLVKFLFYKCDQGDQADPPVAVARDQS 957
Db 943 GLLKQSERVLNNVYIHLGDEDPVRHVAAASLIRLVKFLFYKCDQGDQADPPVAVARDQS 1002
QY 958 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1017
Db 1003 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELLITSTT 1062
QY 1018 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1077
Db 1063 RALTFGCEALCLLSTAFFPVCINSLGWHGCVPLPSASDESRSKCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEENPAATKQEEYVWPALGDRALVPMV 1137
Db 1123 FPLDLSAHQDALIIAGNLLAASAPKSLRSSWASSEENPAATKQEEYVWPALGDRALVPMV 1182
QY 1138 EQLFSLHLLKVNINICAHVLDVAVCPAIPAALPSLTNPPSLSPIRRKKEKEPEGEQASVPL 1197
Db 1183 EQLFSLHLLKVNINICAHVLDVAVCPAIPAALPSLTNPPSLSPIRRKKEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASARSQSDTSGPVTTTSKSSSLGFSYHLPYSLKLDHVLKATHANYKVTLDLQN 1257
Db 1243 SPKKGSEASARSQSDTSGPVTTTSKSSSLGFSYHLPYSLKLDHVLKATHANYKVTLDLQN 1302
QY 1258 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1317
Db 1303 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLYKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPKSKSGRAORLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
Db 1363 LFGTNLASQFDGLSSNPKSKSGRAORLGSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388
Db 1423 VQAEQENDTSG 1433
```


QY 298 LVPVDEHSTLLILGVLTLRLVLPLOQVQKDTSLKSGFVTRKEMEYSPSAEOLVQVY 357
DB 343 LVPVDEHSTLLILGVLTLRLVLPLOQVQKDTSLKSGFVTRKEMEYSPSAEOLVQVY 402
QY 358 ELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSG 417
DB 403 ELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAKEESGGRSG 462
QY 418 SIVELIAGGSSCSPVLSRKQKGVLLGEEAELEDDSESDVSSSALTASVKDEISGEL 477
DB 463 SIVELIAGGSSCSPVLSRKQKGVLLGEEAELEDDSESDVSSSALTASVKDEISGEL 522
QY 478 AASSGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHSSQVVS 537
DB 523 AASSGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHSSQVVS 582
QY 538 AVPSDPAAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNQYLGQIGQP 597
DB 583 AVPSDPAAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNQYLGQIGQP 642
QY 598 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEATEPGD 657
DB 643 QDEDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEATEPGD 702
QY 658 QENKPCRIGKIDGOSTDDDSAPLVHCVLRLLSASFLTGTGKNVLPDRDVRVSVKALALSC 717
DB 703 QENKPCRIGKIDGOSTDDDSAPLVHCVLRLLSASFLTGTGKNVLPDRDVRVSVKALALSC 762
QY 718 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVRGATAILCGTFLICS 777
DB 763 VGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVRGATAILCGTFLICS 822
QY 778 ILRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
DB 823 ILRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 882
QY 838 SSSYSELGLQIIDLVTLRNSSYWLVRTELLETAEIDFRLVSFLKAEKAEHLRGAAHYT 897
DB 883 SSSYSELGLQIIDLVTLRNSSYWLVRTELLETAEIDFRLVSFLKAEKAEHLRGAAHYT 942
QY 898 GLLKQERLVNNVYIHLGDEDPVRHVAAASLIRLVKLFYKCDQOQADPVVAVARDOS 957
DB 943 GLLKQERLVNNVYIHLGDEDPVRHVAAASLIRLVKLFYKCDQOQADPVVAVARDOS 1002
QY 958 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
DB 1003 SVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHELITSTT 1062
QY 1018 RALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILTLLSSAW 1077
DB 1063 RALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILTLLSSAW 1122
QY 1078 FPLDLSAQHODALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGRALVPMV 1137
DB 1123 FPLDLSAQHODALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWPALGRALVPMV 1182
QY 1138 EQLFSLHLKLVINICAHVLDVAPGPAIKAAALPSLTNPSPSLPIRRKKEKEPEGEQASVPL 1197
DB 1183 EQLFSLHLKLVINICAHVLDVAPGPAIKAAALPSLTNPSPSLPIRRKKEKEPEGEQASVPL 1242
QY 1198 SPKKGSEASASROSDTSGPVTTSKSSSLGSFYHLPYSYLLKLDVLCATHANYKVTLDLQ 1257
DB 1243 SPKKGSEASASROSDTSGPVTTSKSSSLGSFYHLPYSYLLKLDVLCATHANYKVTLDLQ 1302
QY 1258 STEFGGFLRSALDVLSQILELATLQDITGKCEVEEILYKSCFSREPMMATVCVQQLLKT 1317
DB 1303 STEFGGFLRSALDVLSQILELATLQDITGKCEVEEILYKSCFSREPMMATVCVQQLLKT 1362
QY 1318 LFGTNLASQFDGLSSNPSKSGOGRQIRLSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1377
DB 1363 LFGTNLASQFDGLSSNPSKSGOGRQIRLSSSVRPGLYHYCFMAYPTHFTQALADASLRNM 1422
QY 1378 VQAEQENDTSG 1388

DB 1423 VQAEQENDTSG 1433
RESULT 9
US-09-724-676-51192
; Sequence 51192, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51192
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51192
Query Match 76.1%; Score 6008; DB 5; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 353 LVQVYELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAKEESGG 412
DB 14 LFQVYELTLHTHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAKEESGG 73
QY 413 RSRSSIVELIAGGSSCSPVLSRKQKGVLLGEEAELEDDSESDVSSSALTASVKDE 472
DB 74 RSRSSIVELIAGGSSCSPVLSRKQKGVLLGEEAELEDDSESDVSSSALTASVKDE 133
QY 473 ISGELAAASGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHS 532
DB 134 ISGELAAASGVSTPGSAGHDITTEQPSOHTLQADSVDLASCDLTSSATDGEDDILSHS 193
QY 533 SSOVSAPVSDPAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNQYLG 592
DB 194 SSOVSAPVSDPAMDNDGTQASSPISDSSQTTTEGPDSAVTPSDSSEIVLDGTDNQYLG 253
QY 593 QLGQPDQDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEA 652
DB 254 QLGQPDQDEEATGILPDEASEAFRNSSMALQQAHLKNMSHCRQPSDSSVDKXVLRDEA 313
QY 653 TEPGQENKPCRIGKIDGOSTDDDSAPLVHCVLRLLSASFLTGTGKNVLPDRDVRVSVKA 712
DB 314 TEPGQENKPCRIGKIDGOSTDDDSAPLVHCVLRLLSASFLTGTGKNVLPDRDVRVSVKA 373
QY 713 LALSCVGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVRGATAILCG 772
DB 374 LALSCVGAVALHPESFFSKLYKVPDTPTEPEQYVSDILNYIDHGPQVRGATAILCG 433
QY 773 TLICSTLSRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC 832
DB 434 TLICSTLSRSRPHVGDWMTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC 493
QY 833 VMSLCSYSELGLQIIDLVTLRNSSYWLVRTELLETAEIDFRLVSFLKAEKAEHLRG 892
DB 494 VMSLCSYSELGLQIIDLVTLRNSSYWLVRTELLETAEIDFRLVSFLKAEKAEHLRG 553
QY 893 AHHTYGLLKQERLVNNVYIHLGDEDPVRHVAAASLIRLVKLFYKCDQOQADPVVAV 952
DB 554 AHHTYGLLKQERLVNNVYIHLGDEDPVRHVAAASLIRLVKLFYKCDQOQADPVVAV 613
QY 953 ARDOSSVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHEL 1012
DB 614 ARDOSSVYLLKLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMNNLSRVIAAVSHEL 673
QY 1013 ITSTTRALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 1072
DB 674 ITSTTRALTFGCCEALCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 733


```

QY 1073 LSSAWFPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEEEANPAATKOEWWPALGDRA 1132
DQ 734 LSSAWFPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEEEANPAATKOEWWPALGDRA 793
QY 1133 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPALTNPSPSPIRRKGKEKEPGEQ 1192
DQ 794 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPALTNPSPSPIRRKGKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASQSDTSQPVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 1252
DQ 854 ASVPLSPKKGSEASASQSDTSQPVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 1312
DQ 914 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1033
DQ 974 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSGWFDVLQKVSTQTKNLTSVTKNRADKNATHNIRLFEPLVICAL 1432
DQ 1034 SLRNMVQAEQENDTSGWFDVLQKVSTQTKNLTSVTKNRADKNATHNIRLFEPLVICAL 1093
QY 1433 KOYTTTTCVQKQVLDLQALVQLRVNYCLDSDQVFIGFVLKQFEYIEVQFSEAI 1492
DQ 1094 KOYTTTTCVQKQVLDLQALVQLRVNYCLDSDQVFIGFVLKQFEYIEVQFSEAI 1153
QY 1493 IPNIFFLVLLSYERYHSHQIIGIPKIIQICDGMASGRKA 1533
DQ 1154 IPNIFFLVLLSYERYHSHQIIGIPKIIQICDGMASGRKA 1194

RESULT 10
US-09-724-676A-51192
; Sequence 51192, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; NUMBER OF SEQ ID NOS: 2000-11-28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51192
; LENGTH: 1763
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51192

Query Match 76.1%; Score 6008; DB 5; Length 1763;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 353 LVQVVELTLLHQTQHDHNVVTGALLELQQLFRTPPPELLQTLTAVGIGQLTAKESEGG 412
DQ 14 LVQVVELTLLHQTQHDHNVVTGALLELQQLFRTPPPELLQTLTAVGIGQLTAKESEGG 73
QY 413 RSRSGSIVELIAGGSSCSPVLRSKQKGVLLGEEALEDDSESRSDVSSSALTASVKDE 472
DQ 74 RSRSGSIVELIAGGSSCSPVLRSKQKGVLLGEEALEDDSESRSDVSSSALTASVKDE 133
QY 473 ISGELASSGVSTPGSAGHDIIITEQPSQHTLQADSVDLASCDLTSSATDGEDILSHS 532
DQ 134 ISGELASSGVSTPGSAGHDIIITEQPSQHTLQADSVDLASCDLTSSATDGEDILSHS 193
QY 533 SSQVSAYVSPDAMDNDGTQASSPISDSQTTTEGPDOSAVTPDSSSEIVLDGTDNOYGL 592
DQ 194 SSQVSAYVSPDAMDNDGTQASSPISDSQTTTEGPDOSAVTPDSSSEIVLDGTDNOYGL 253
QY 593 QIGQPODEDEATGILPDEASEAFRNSMALQQAHLKNNSHCHRCPSDSSVDKFLRDEA 652
DQ 254 QIGQPODEDEATGILPDEASEAFRNSMALQQAHLKNNSHCHRCPSDSSVDKFLRDEA 313

```

```

QY 653 TEPQDQENKPCRIKGDITGQSTDDDSAPLVHCVRLLSASFLLTGCKNVLVDPDROVRVSKA 712
DQ 314 TEPQDQENKPCRIKGDITGQSTDDDSAPLVHCVRLLSASFLLTGCKNVLVDPDROVRVSKA 373
QY 713 LALSCVGAVALHPESFESKLYKVPDLDTTEYPERQYYSIDLINYLIDHGDPQVGRATILCG 772
DQ 374 LALSCVGAVALHPESFESKLYKVPDLDTTEYPERQYYSIDLINYLIDHGDPQVGRATILCG 433
QY 773 TLICSIILSRFRHFHGWDMGTIRTLTGNTFSLADCIPLLRKTKLDESSVTCKLACTAVRNC 832
DQ 434 TLICSIILSRFRHFHGWDMGTIRTLTGNTFSLADCIPLLRKTKLDESSVTCKLACTAVRNC 493
QY 833 VMSLCSSSYSELGLQLIIDVLTLRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLHRG 892
DQ 494 VMSLCSSSYSELGLQLIIDVLTLRNSSYWLVRTELLETLAEIDFRLVSLFLEAKAENLHRG 553
QY 893 AHHTYGLLKLQERVLNNVTHLLGDEDPVRHVAASLIRLVPKLFYKCDQGOADPVAV 952
DQ 554 AHHTYGLLKLQERVLNNVTHLLGDEDPVRHVAASLIRLVPKLFYKCDQGOADPVAV 613
QY 953 ARQSSVYLKLLMHETOPPSHFVSITIRYGYNLLPSITDVTMENNLSRVIAAVSHEL 1012
DQ 614 ARQSSVYLKLLMHETOPPSHFVSITIRYGYNLLPSITDVTMENNLSRVIAAVSHEL 673
QY 1013 ITSTTRALTFCCEALCLLSTAPVCIWSLGHGCVPPPLSASDESRSKSCVTGMATMILT 1072
DQ 674 ITSTTRALTFCCEALCLLSTAPVCIWSLGHGCVPPPLSASDESRSKSCVTGMATMILT 733
QY 1073 LSSAWFPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEEEANPAATKOEWWPALGDRA 1132
DQ 734 LSSAWFPLDLSAHQDALILAGNLLAASAPKSLRSSWASSEEEANPAATKOEWWPALGDRA 793
QY 1133 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPALTNPSPSPIRRKGKEKEPGEQ 1192
DQ 794 LVPWVEQLFSLHLLKYNINCAHVDVAVGPAKALPALTNPSPSPIRRKGKEKEPGEQ 853
QY 1193 ASVPLSPKKGSEASASQSDTSQPVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 1252
DQ 854 ASVPLSPKKGSEASASQSDTSQPVTTSSKSSLGSEFVHLPSYKLHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 1312
DQ 914 LDQNSTEKGFLRSALDVLISQILELATLQDQKCVBEIILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1372
DQ 974 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSGWFDVLQKVSTQTKNLTSVTKNRADKNATHNIRLFEPLVICAL 1432
DQ 1034 SLRNMVQAEQENDTSGWFDVLQKVSTQTKNLTSVTKNRADKNATHNIRLFEPLVICAL 1093
QY 1433 KOYTTTTCVQKQVLDLQALVQLRVNYCLDSDQVFIGFVLKQFEYIEVQFSEAI 1492
DQ 1094 KOYTTTTCVQKQVLDLQALVQLRVNYCLDSDQVFIGFVLKQFEYIEVQFSEAI 1153
QY 1493 IPNIFFLVLLSYERYHSHQIIGIPKIIQICDGMASGRKA 1533
DQ 1154 IPNIFFLVLLSYERYHSHQIIGIPKIIQICDGMASGRKA 1194

RESULT 11
US-09-724-676-51191
; Sequence 51191, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```


; SEQ ID NO 51191
; LENGTH: 1940
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51191

Query Match 76.1%; Score 6008; DB 5; Length 1940;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 353 LVQVYELTTHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 412
Db 14 LFQVYELTTHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 73
QY 413 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVDSVDFVLRDEA 472
Db 74 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVDSVDSVDFVLRDEA 133
QY 473 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEDDILSHS 532
Db 134 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEDDILSHS 193
QY 533 SSOVSAPVSDPAMDNDGTQASSPISDSSTQTTTEGPDSAVTPSDSSEIVLDGTNOYLGL 592
Db 194 SSOVSAPVSDPAMDNDGTQASSPISDSSTQTTTEGPDSAVTPSDSSEIVLDGTNOYLGL 253
QY 593 QIQOQODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLVRDEA 652
Db 254 QIQOQODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLVRDEA 313
QY 653 TEPGQDNKPCRIGDQIGOSTDDSDAPLVHCVRLLSASFLLTGKNNLVDPDRVSVKA 712
Db 314 TEPGQDNKPCRIGDQIGOSTDDSDAPLVHCVRLLSASFLLTGKNNLVDPDRVSVKA 613
QY 953 ARQSSVYLKLLMHETQPPSHFSVTITRIYGNLLPSITDVTMNNLSRVIAAVSHSL 1012
Db 614 ARQSSVYLKLLMHETQPPSHFSVTITRIYGNLLPSITDVTMNNLSRVIAAVSHSL 673
QY 1013 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILTL 1072
Db 674 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILTL 733
QY 1073 LSSANFPDLISAQDALILAGNLLASAPKSLRSSWASBEANPAATKQEEVWPALGDRA 1132
Db 734 LSSANFPDLISAQDALILAGNLLASAPKSLRSSWASBEANPAATKQEEVWPALGDRA 793
QY 1133 LVPVVEQLFHLKLVINICAHVLDVDPAGPAIKAAPLSITNPPSLPIRRKKEKEPEGQ 1192
Db 794 LVPVVEQLFHLKLVINICAHVLDVDPAGPAIKAAPLSITNPPSLPIRRKKEKEPEGQ 853
QY 1193 ASVPLSPKKGEASAAASRQSDTSGPVVTTSSSSLSGSGFYHLPSYKLHDLVKATHANYKVT 1252
Db 854 ASVPLSPKKGEASAAASRQSDTSGPVVTTSSSSLSGSGFYHLPSYKLHDLVKATHANYKVT 913
QY 1253 LDQNSTEKGFGFLRSALDVLSQLILELATLQDYGKVEETLQDYGKSCFREPMMATVCVQ 1312
Db 914 LDQNSTEKGFGFLRSALDVLSQLILELATLQDYGKVEETLQDYGKSCFREPMMATVCVQ 973
```

```
QY 1313 QLLKTLFCTNLASQFDGLSSNPSSKSGRAORLGLSSSVRPGLYHYCFMAPYTHFTQALADA 1372
Db 974 QLLKTLFCTNLASQFDGLSSNPSSKSGRAORLGLSSSVRPGLYHYCFMAPYTHFTQALADA 1033
QY 1373 SLRNNWAOEONDTSQWFDVLQKVSTOLKTNLTSVTKNRADKNAIHHIRLFPPLVIKAL 1432
Db 1034 SLRNNWAOEONDTSQWFDVLQKVSTOLKTNLTSVTKNRADKNAIHHIRLFPPLVIKAL 1093
QY 1433 KOYTTTTCVOLQKQVLDLALQVLQVRVNYCCLLSDQVFIGFVLKQFYEIVGQFRESEAI 1492
Db 1094 KOYTTTTCVOLQKQVLDLALQVLQVRVNYCCLLSDQVFIGFVLKQFYEIVGQFRESEAI 1153
QY 1493 IPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1154 IPNIFFFVLVLSYERYHSKQIIGIPKIIQLCDGIMASGRKA 1194
```

RESULT 12

US-09-724-676A-51191

; Sequence 51191, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724, 676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 51191

; LENGTH: 1940

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-51191

Query Match 76.1%; Score 6008; DB 5; Length 1940;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 353 LVQVYELTTHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 412
Db 14 LFQVYELTTHHTQHODHNVTGALQLLQQLFRTPPPELLQTLTAVGGIGQLTAAKEESGG 73
QY 413 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVDSVDSVDFVLRDEA 472
Db 74 RSRGSGIVELIAGGSSCPVLSRKQKGVLLGEEAELEDDSESRSVDSVDSVDFVLRDEA 133
QY 473 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEDDILSHS 532
Db 134 ISGELAASSGVSTPGSAGHDITTEQPRSOHTLQADSVDLASCDLTSSATDGEDDILSHS 193
QY 533 SSOVSAPVSDPAMDNDGTQASSPISDSSTQTTTEGPDSAVTPSDSSEIVLDGTNOYLGL 592
Db 194 SSOVSAPVSDPAMDNDGTQASSPISDSSTQTTTEGPDSAVTPSDSSEIVLDGTNOYLGL 253
QY 593 QIQOQODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLVRDEA 652
Db 254 QIQOQODEEATGILPDEASEAFRNSMALQOAHLLKNMHCROPSDSSVDKFLVRDEA 313
QY 653 TEPGQDNKPCRIGDQIGOSTDDSDAPLVHCVRLLSASFLLTGKNNLVDPDRVSVKA 712
Db 314 TEPGQDNKPCRIGDQIGOSTDDSDAPLVHCVRLLSASFLLTGKNNLVDPDRVSVKA 373
QY 713 LALSCVGAVALHPESFFSKLYKVPDLDTTEYPEEQVSDILNYIDHGPQVARGATAILCG 772
Db 374 LALSCVGAVALHPESFFSKLYKVPDLDTTEYPEEQVSDILNYIDHGPQVARGATAILCG 433
QY 773 TLICSIILSRFRHVGDMGTITRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 832
Db 434 TLICSIILSRFRHVGDMGTITRTLTGNTFSLADCIPLLRKTLKDESSVTCCKLACTAVRNC 493
QY 833 VMSLCSSSYSELGLOLIDVLTLRNSSYLWVTRTELETLAEIDPRLVSLFLEAKAENLHRG 892
```



```

Db 494 VMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELLETAEIDPRLVSFLEAKAENLHRG 553
QY 893 AHHTYGLKQLQERVLNNVVIHLLGDEDRVRHVAAASLIRLVPKLFYKCDGQADPPVAV 952
Db 554 AHHTYGLKQLQERVLNNVVIHLLGDEDRVRHVAAASLIRLVPKLFYKCDGQADPPVAV 613
QY 953 ARDQSSVYLKLLMHETQPPSHFSVSSTIIRYGYNLLPSITDVTMNNLSRVIAAVSHL 1012
Db 614 ARDQSSVYLKLLMHETQPPSHFSVSSTIIRYGYNLLPSITDVTMNNLSRVIAAVSHL 673
QY 1013 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILT 1072
Db 674 ITSTTRALTFCGCEALCLLSTAFPVCIVSLGWHCGVPPPLSASDESRSKCTVGMATMILT 733
QY 1073 LSSANFPDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWALGDRA 1132
Db 734 LSSANFPDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEEVWALGDRA 793
QY 1133 LVPVMEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSLSPIRRKKEKEPEQ 1192
Db 794 LVPVMEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSLSPIRRKKEKEPEQ 853
QY 1193 ASVPLSPKKGSSASASRSQSDTSGPVTTSKSSLSGSFYHLPYSYKLHDLVLAATHANYKVT 1252
Db 854 ASVPLSPKKGSSASASRSQSDTSGPVTTSKSSLSGSFYHLPYSYKLHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKGFGFLRSALDVLQIILELATLQDQKVEEILGYLKSCFSREPMATVCVQ 1312
Db 914 LDQNSTEKGFGFLRSALDVLQIILELATLQDQKVEEILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFTCNLASQFDGLSSNPSKSGRAQLRGSSSVRPGLYHYCFMAYPTHFTQALADA 1372
Db 974 QLLKTLFTCNLASQFDGLSSNPSKSGRAQLRGSSSVRPGLYHYCFMAYPTHFTQALADA 1033
QY 1373 SLRNVVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNAIHNIRLFEPLVIKAL 1432
Db 1034 SLRNVVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNAIHNIRLFEPLVIKAL 1093
QY 1433 QYTTTTTCVQKQVLDLQALQVLRVNYCLLSDSDQVFIGVLFKQFIEVQFRESFAI 1492
Db 1094 QYTTTTTCVQKQVLDLQALQVLRVNYCLLSDSDQVFIGVLFKQFIEVQFRESFAI 1153
QY 1493 IPNIFFLVLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1154 IPNIFFLVLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1194

RESULT 13
US-09-724-676-51190
; Sequence 51190, Application us/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51190
; LENGTH: 2690
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-51190

Query Match 69.5%; Score 5484; DB 5; Length 2690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 DVSSSALTASVKDEISGELAASSGVSTPGSAGHDIIITEQPSQHTLOADSVDLASCDLTS 518
Db 8 DVSSSALTASVKDEISGELAASSGVSTPGSAGHDIIITEQPSQHTLOADSVDLASCDLTS 67
QY 519 SATDGEDEEDILSHSSQVSAVPSDPAMDNDGTQASSPISDSSQTTTEGPDPSAVTPSDSS 578

```

```

Db 68 SATDGEDEEDILSHSSQVSAVPSDPAMDNDGTQASSPISDSSQTTTEGPDPSAVTPSDSS 127
QY 579 EIVLDGTDNOYLGIGIQPODEDEATGILPDEASEAFRNSSMALQQAHLKNNSHCRQP 638
Db 128 EIVLDGTDNOYLGIGIQPODEDEATGILPDEASEAFRNSSMALQQAHLKNNSHCRQP 187
QY 639 SDSSVDKFLRDEATEPGDQENPCRTKGDIGQSTDDDSAPLVHCVRLLSASFLLTGCKN 698
Db 188 SDSSVDKFLRDEATEPGDQENPCRTKGDIGQSTDDDSAPLVHCVRLLSASFLLTGCKN 247
QY 699 VLPDRDVRVSVKALASCVAALHPESFESKLYKVPDPTTEYPEQVVSIDLNVYIDH 758
Db 248 VLPDRDVRVSVKALASCVAALHPESFESKLYKVPDPTTEYPEQVVSIDLNVYIDH 307
QY 759 GDPQVRGATAILCGTLLCSILSRSHRPHVGDWMTGIRTLTGNTFSLADICPLLRKTKDES 818
Db 308 GDPQVRGATAILCGTLLCSILSRSHRPHVGDWMTGIRTLTGNTFSLADICPLLRKTKDES 367
QY 819 SVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELLETAEIDFRL 878
Db 368 SVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVLTNRSSYWLVRTELLETAEIDFRL 427
QY 879 VSFLEAKAENLHRGAHHYTGLLKQERVLNNVVIHLLGDEDPVRRHVAAASLIRLVPKLF 938
Db 428 VSFLEAKAENLHRGAHHYTGLLKQERVLNNVVIHLLGDEDPVRRHVAAASLIRLVPKLF 487
QY 939 KYCQDQADPVAVARDQSSVYLKLLMHETQPPSHFSVSSTIIRYGYNLLPSITDVTME 998
Db 488 KYCQDQADPVAVARDQSSVYLKLLMHETQPPSHFSVSSTIIRYGYNLLPSITDVTME 547
QY 999 NNLSRVTAAYSHELITSTTRALTFCGCEALCLLSTAPPVCIVSLGWHCGVPPPLSASDES 1058
Db 548 NNLSRVTAAYSHELITSTTRALTFCGCEALCLLSTAPPVCIVSLGWHCGVPPPLSASDES 607
QY 1059 KSCITVGWATMILTLLSSANFPDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPA 1118
Db 608 KSCITVGWATMILTLLSSANFPDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPA 667
QY 1119 TKQEEVWALGDRAVPMVEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSL 1178
Db 668 TKQEEVWALGDRAVPMVEQLFSLHLLKVINICAHVLDVAVPAPAIKAALPSLTNPPSL 727
QY 1179 PIRKKEKEPEGQASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPYSYKL 1238
Db 728 PIRKKEKEPEGQASVPLSPKKGSEASASRSQSDTSGPVTTSKSSLSGSFYHLPYSYKL 787
QY 1239 HDVLKATHANYKVTLDLQNSTEKGFGFLRSALDVLQIILELATLQDQKVEEILGYLKS 1298
Db 788 HDVLKATHANYKVTLDLQNSTEKGFGFLRSALDVLQIILELATLQDQKVEEILGYLKS 847
QY 1299 CFSREPMMATVCVOQLLKTFLGNTLASQFDGLSSNPSKSGRAQLRGSSSVRPGLYHYCF 1358
Db 848 CFSREPMMATVCVOQLLKTFLGNTLASQFDGLSSNPSKSGRAQLRGSSSVRPGLYHYCF 907
QY 1359 MAPYTHFTQALADASLRNVVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNAI 1418
Db 908 MAPYTHFTQALADASLRNVVQAEQNDTSGWFDVLQKVSQTLKNTLTSVTKNRADKNAI 967
QY 1419 NHIRLFEPLVIKALQYTTTTTCVQKQVLDLQALQVLRVNYCLLSDSDQVFIGVLFKQF 1478
Db 968 NHIRLFEPLVIKALQYTTTTTCVQKQVLDLQALQVLRVNYCLLSDSDQVFIGVLFKQF 1027
QY 1479 EYIEVQFRESFAIIPNIFFLVLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1533
Db 1028 EYIEVQFRESFAIIPNIFFLVLSYERYHSHKQIIGIPKIIQLCDGIMASGRKA 1082

RESULT 14
US-09-724-676A-51190
; Sequence 51190, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD

```



```

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 123181.4 Compugen
; CURRENT APPLICATION NUMBER: US709/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51190
; LENGTH: 2690
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-724-676A-51190

```

```
Query Match      69.5%; Score 5484; DB 5; Length 2690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1075; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	459	DVSSSALTASVKDEITSIGELAASSGVSTPGSAGHDIITTEQPRSQHITLOADSVDIASCOLTIS	518
Db	8	DVSSSALTASVKDEITSIGELAASSGVSTPGSAGHDIITTEQPRSQHITLOADSVDIASCOLTIS	67
Qy	519	SATDGBDEEDILSHSSSQSVSAVSPDAMPNDGTOAQSSPIDSSSQTTTEGPDSAVTPSDSS	578
Db	68	SATDGBDEEDILSHSSSQSVSAVSPDAMPNDGTOAQSSPIDSSSQTTTEGPDSAVTPSDSS	127
Qy	579	EIVLDGTDNQYLGQLIQGQPODEDEATGILPDEASEAFNRSSMALQQAHLKNNMSHCROP	638
Db	128	EIVLDGTDNQYLGQLIQGQPODEDEATGILPDEASEAFNRSSMALQQAHLKNNMSHCROP	187
Qy	639	SDSSVDKFVLRDEATEPGDOENKPCRIGDITGQSTDDDSAPLVHCVRLLSASFLLTGGKN	698
Db	188	SDSSVDKFVLRDEATEPGDOENKPCRIGDITGQSTDDDSAPLVHCVRLLSASFLLTGGKN	247
Qy	699	VLVPDRDVRVSKALALSCVGAALVHPESFESKLYKVPDITTEYPEQVXSUILNYIDH	758
Db	248	VLVPDRDVRVSKALALSCVGAALVHPESFESKLYKVPDITTEYPEQVXSUILNYIDH	307
Qy	759	GDPOVRGATAILCGTLICISILSRSFHWGDMGTIRLTGNTPFSLADCIPLLRKTTLKDES	818
Db	308	GDPOVRGATAILCGTLICISILSRSFHWGDMGTIRLTGNTPFSLADCIPLLRKTTLKDES	367
Qy	819	SVTCKLACTAVRNCVMSLCSSSYSELGLOLIDVLTLRNSSYMLVRLTLELTLAEIDFRL	878
Db	368	SVTCKLACTAVRNCVMSLCSSSYSELGLOLIDVLTLRNSSYMLVRLTLELTLAEIDFRL	427
Qy	879	VSFLEAKAENLRGAAHYHTGLLKLQERVLLNNVVTHLLGDEDPVRVRHVAASLRLVPKLF	938
Db	428	VSFLEAKAENLRGAAHYHTGLLKLQERVLLNNVVTHLLGDEDPVRVRHVAASLRLVPKLF	487
Qy	939	YKCDQGOADPVAVARDQSSVYLKLLMHETOPPSHFSVSTITRIYRGYNLLPSTDVTME	998
Db	488	YKCDQGOADPVAVARDQSSVYLKLLMHETOPPSHFSVSTITRIYRGYNLLPSTDVTME	547
Qy	999	NNLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFPVCIIWSLGHWCVPPLSASDES	1058
Db	548	NNLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFPVCIIWSLGHWCVPPLSASDES	607
Qy	1059	KSCTVGMATMILTLSSAWFPDLDSAHODALILAGNLLAASAPKSLRSSWASSEEEANPAA	1118
Db	608	KSCTVGMATMILTLSSAWFPDLDSAHODALILAGNLLAASAPKSLRSSWASSEEEANPAA	667
Qy	1119	TKQEEVWPALGDRLVPWYEQLFSLHLKVINICAHVLDVAPGPAIKAAALPSTINPPSL	1178
Db	668	TKQEEVWPALGDRLVPWYEQLFSLHLKVINICAHVLDVAPGPAIKAAALPSTINPPSL	727
Qy	1179	PIRRKGEKEPGBQASVPLSPKKGSEASARSQSDTSGPVTTSKSSSLGSGFYHLPSYKL	1238
Db	728	PIRRKGEKEPGBQASVPLSPKKGSEASARSQSDTSGPVTTSKSSSLGSGFYHLPSYKL	787
Qy	1239	HDVLKATHANYKVTLDLONSTEFGGFLRSALDVLSQITLELATLQDTCKCVEETLGLYKS	1298
Db	788	HDVLKATHANYKVTLDLONSTEFGGFLRSALDVLSQITLELATLQDTCKCVEETLGLYKS	847
Qy	1299	CF9REPMATVVCVQQLKTLFTGNLASOPDGLSSNPSKSGQRAORLGSSSVRPGCLYHCF	1358

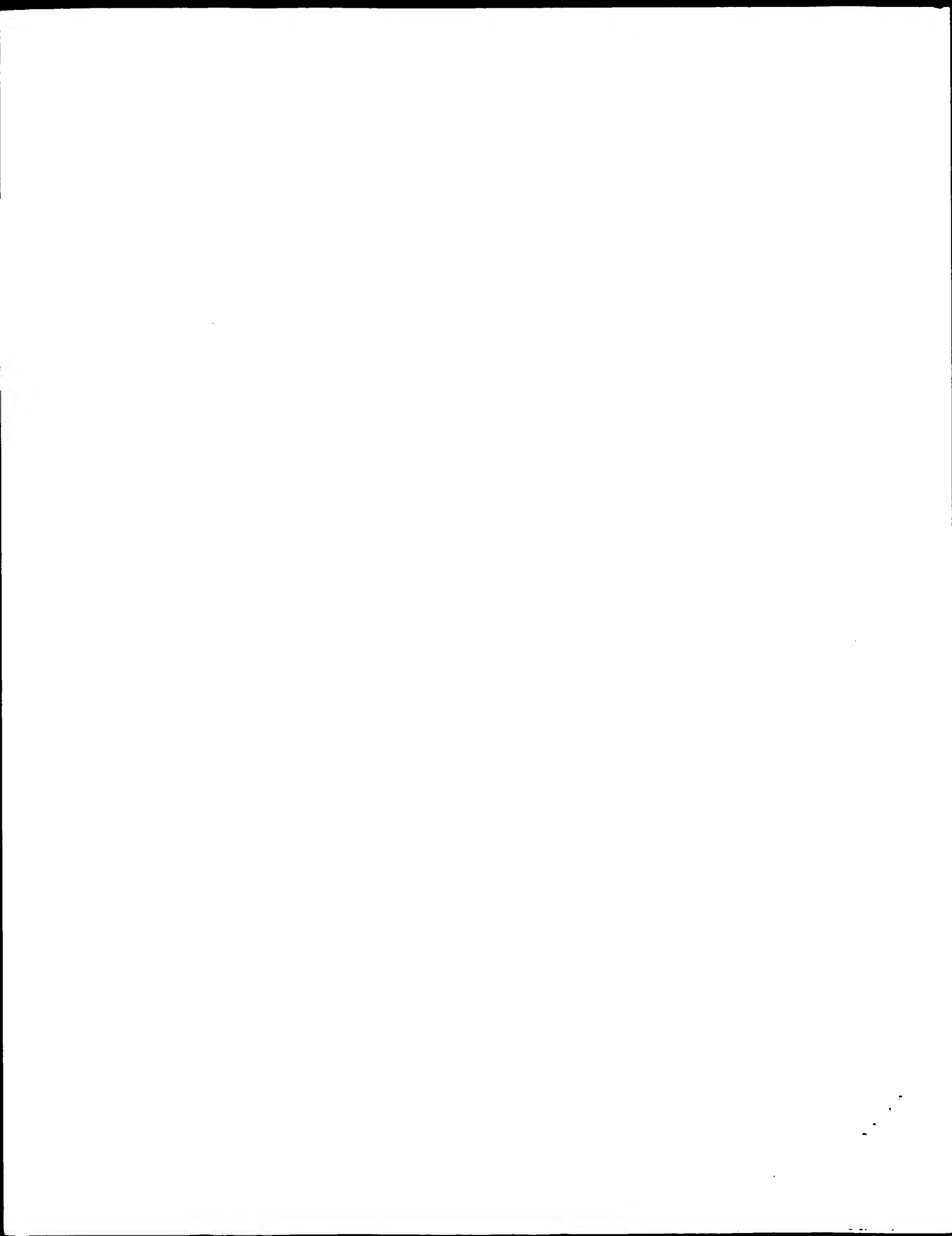
Db	848	CFSEPEPMATVCVQQLKTLFTGNTLASQDFGLSSNPSSQGRAQLGSSSVRPGLYHYCF	907
Qy	1359	MAPYTHFTQALADASLRNNVQAEQENDTSGWPDVLQKVSTOLKNTLTSVTKNRADKNAIH	1418
Db	908	MAPYTHFTQALADASLRNNVQAEQENDTSGWPDVLQKVSTQLKTNLTSTVTKNRADKNAIH	967
Qy	1419	NHRLRLEPLVIKALKOYTTTTCVQKQKQVLDLLAQLVOLRVNYCCLLDSQVFIGFVLKQF	1478
Db	968	NHRLRLEPLVIKALKOYTTTTCVQKQKQVLDLLAQLVOLRVNYCCLLDSQVFIGFVLKQF	1027
Qy	1479	EYIEVQGFRESEAIIPNIFFFVLVLSYERYHSHQIIGIPKIIQLCDGIMASGRKA	1533
Db	1028	EYIEVQGFRESEAIIPNIFFFVLVLSYERYHSHQIIGIPKIIQLCDGIMASGRKA	1082
RESULT 15			
US-09-724-676-51181			
; Sequence 51181, Application US/09724676			
; GENERAL INFORMATION:			
; APPLICANT: Comugen LTD			
; TITLE OF INVENTION: Variants of alternative splicing			
; FILE REFERENCE: 129181.4 Comugen			
; CURRENT APPLICATION NUMBER: US/09/724,676			
; CURRENT FILING DATE: 2000-11-28			
; NUMBER OF SEQ ID NOS: 97222			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 51181			
; LENGTH: 1049			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-724-676-51181			

Query Match	66.8%	Score 5273;	DB 5;	Length 1049;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1035: Conservative	0;	Mismatches	1;	Indels 0;
Gaps	0;			

Qy	353	LVQYVELTHHTHQDHNVVTGALELLQQLFRTPPPELLQTTAVGGIQLTAAKEESGG	412
Db	14	LFQYVELTHHTHQDHNVVTGALELLQQLFRTPPPELLQTTAVGGIQLTAAKEESGG	73
Qy	413	RRSGSIVELIAGGSGSPVLRSRKQKGVLLGEBEALDDSESRSNDSSSALTASVXDE	472
Db	74	RRSGSIVELIAGGSGSPVLRSRKQKGVLLGEBEALDDSESRSNDSSSALTASVXDE	133
Qy	473	ISGELAASSGVSTPGSAGHDIIITEQPRSQHTLQADSVDLASCDLTSSATDGOEBEILSHS	533
Db	134	ISGELAASSGVSTPGSAGHDIIITEQPRSQHTLQADSVDLASCDLTSSATDGOEBEILSHS	193
Qy	533	SSQYSAVPSDPMDLNDGTOACSSPIDSSQTTTEGPDSAVTPSDSEIIVLDCDNOYLCL	592
Db	194	SSQYSAVPSDPMDLNDGTOACSSPIDSSQTTTEGPDSAVTPSDSEIIVLDCDNOYLCL	253
Qy	593	QIGQPQDEDEATGILPDEASEAFNRSSMALQOAHLLKNMHCROPSDSSVDKPFVLRDEA	652
Db	254	QIGQPQDEDEATGILPDEASEAFNRSSMALQOAHLLKNMHCROPSDSSVDKPFVLRDEA	313
Qy	653	TEPGQENKPCRIGDIIQGSTDDDSAPLVHCYRLLSASFLLTGKKNLVDPDRDVRVSVA	712
Db	314	TEPGQENKPCRIGDIIQGSTDDDSAPLVHCYRLLSASFLLTGKKNLVDPDRDVRVSVA	373
Qy	713	LALSCVGAVALHPESFFSKLYKVPJDTTETPEEQYVSDILNYIDHGDPPQVRGATAILCG	772
Db	374	LALSCVGAVALHPESFFSKLYKVPJDTTETPEEQYVSDILNYIDHGDPPQVRGATAILCG	433
Qy	773	TLICSIILSRFRHVGDMWGIIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC	832
Db	434	TLICSIILSRFRHVGDMWGIIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNC	493
Qy	833	VMSLCSSSYSELGLQLIIDLVTLRNSSYWLVRTELETETLAEIDFRULVSFLEAKENLHRG	892
Db	494	VMSLCSSSYSELGLQLIIDLVTLRNSSYWLVRTELETETLAEIDFRULVSFLEAKENLHRG	553

QY 893 AHHTGLLKQERVLNNVVIHLGDEDPVRHVAASLIRLPKLFYKCDQOQADPVVAV 952
Db 554 AHHTGLLKQERVLNNVVIHLGDEDPVRHVAASLIRLPKLFYKCDQOQADPVVAV 613
QY 953 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHEL 1012
Db 614 ARDOSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHEL 673
QY 1013 ITSTTRALTFGCEALCCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 1072
Db 674 ITSTTRALTFGCEALCCLLSTAFFVCINSLGWHCGVPPPLSASDESRSKCTVGMATMILT 733
QY 1073 LSSANFPDLQSAHQDALIIDAGNLLAASAPKSLRSSWASSEEEANPAATKQEEVWPALGDRA 1132
Db 734 LSSANFPDLQSAHQDALIIDAGNLLAASAPKSLRSSWASSEEEANPAATKQEEVWPALGDRA 793
QY 1133 LVPMEQOLFESHLKVINICAHVLDVAPGPAIKAALPSLTNPSPSIRRRKKEKEPEGEQ 1192
Db 794 LVPMEQOLFESHLKVINICAHVLDVAPGPAIKAALPSLTNPSPSIRRRKKEKEPEGEQ 853
QY 1193 ASVPLSPKKGSEASASROSQDTSQPVTTSSSLGSGFYHLPKLYLKHDLVLAATHANYKVT 1252
Db 854 ASVPLSPKKGSEASASROSQDTSQPVTTSSSLGSGFYHLPKLYLKHDLVLAATHANYKVT 913
QY 1253 LDQNSTEKFGFLRSALDVLQILELATLQDIGKCVBEILGYLKSCFSREPMATVCVQ 1312
Db 914 LDQNSTEKFGFLRSALDVLQILELATLQDIGKCVBEILGYLKSCFSREPMATVCVQ 973
QY 1313 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRGLGSSSVRPGLYHYCFMAYTHFTQALADA 1372
Db 974 QLLKTLFGTNLASQFDGLSSNPSKSGRAQRGLGSSSVRPGLYHYCFMAYTHFTQALADA 1033
QY 1373 SLRNMVQAEQENDTSG 1388
Db 1034 SLRNMVQAEQENDTSG 1049

Search completed: January 28, 2003, 14:39:40
Job time : 63.0972 secs



QY 238 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 297
Db 241 FGNFANDNEIKVLLKAFIANLKSSPTIRRTAAGSAVSICQHSRRTOYFYSWLLNVLLGL 300
QY 298 LVPVEDEHSTLLGLVLLRLYLVLPLLOQOVKDTSLKGSFGVTRKEMEYSPSAEQLVQVY 357
Db 301 LVPVEDEHSTLLGLVLLRLYLVLPLLOQOVKDTSLKGSFGVTRKEMEYSPSAEQLVQVY 360
QY 358 ELTLHHTQHDHNVVTALELLQQLFRTPPPELLOLTAAGGIGQLTAAKESGGRSRG 417
Db 361 ELTLHHTQHDHNVVTALELLQQLFRTPPPELLOLTAAGGIGQLTAAKESGGRSRG 420
QY 418 SIVELIAGGSGSPVLSRKQKGVLLGEEFEALDDSDSDVSSSALTASVKDEISGBL 477
Db 421 SIVELIAGGSGSPVLSRKQKGVLLGEEFEALDDSDSDVSSSALTASVKDEISGBL 480
QY 478 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATGDDEEDILSHSSQVS 537
Db 481 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATGDDEEDILSHSSQVS 540
QY 538 AVPSDPAMDNDGTOASSPISDSOTTTEGPDPSAVTPSPDSSEIVLDGTDNQVGLQIQGP 597
Db 541 AVPSDPAMDNDGTOASSPISDSOTTTEGPDPSAVTPSPDSSEIVLDGTDNQVGLQIQGP 600
QY 598 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVLRDEATEPGD 657
Db 601 QDEDEATGILPDEASEAFRNSSMALQQAHLKNNKSHCRQPSDSSVDKFLVLRDEATEPGD 660
QY 658 QENKPCRIKGDIGOSTDDSDAPLVHCVRLLSASFLITGCKNVLPDRDVRVSVKALALSC 717
Db 661 QENKPCRIKGDIGOSTDDSDAPLVHCVRLLSASFLITGCKNVLPDRDVRVSVKALALSC 720
QY 718 VGAVALHPESFSLKYKVPLOTTTEPEQYVSDILNYIDHGDPOVGRGATATLTCGLTICS 777
Db 721 VGAVALHPESFSLKYKVPLOTTTEPEQYVSDILNYIDHGDPOVGRGATATLTCGLTICS 780
QY 778 ILSRFRHVGDMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
Db 781 ILSRFRHVGDMGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQIIDLVLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLHRCAGHYT 897
Db 841 SSSYSELGLQIIDLVLTRNSSYWLVRTELLETLAEIDFLVSLFLEAKAENLHRCAGHYT 900
QY 898 GLLKQERYLVNNAIHLGDEDPVRRHVAASLIIRLPKFLYKCDQGOADPVAVARQDS 957
Db 901 GLLKQERYLVNNAIHLGDEDPVRRHVAASLIIRLPKFLYKCDQGOADPVAVARQDS 960
QY 958 SYVLKLLMHETOPPSHFVSSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTT 1017
Db 961 SYVLKLLMHETOPPSHFVSSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTT 1020
QY 1018 RALTFCCEALCLLTAPVPCVTLWSLGHGCVPPPLSASDESRSKCTVGMATWILLILLSAW 1077
Db 1021 RALTFCCEALCLLTAPVPCVTLWSLGHGCVPPPLSASDESRSKCTVGMATWILLILLSAW 1080
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEVNPAALGDALVPMV 1137
Db 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKQEVNPAALGDALVPMV 1140
QY 1138 EQLFSHLLKVINICARHVLDDVAPGPAIKAALPSLTNPSPSLPIRRKKEKEPGEQASVPL 1197
Db 1141 EQLFSHLLKVINICARHVLDDVAPGPAIKAALPSLTNPSPSLPIRRKKEKEPGEQASVPL 1200
QY 1198 SPKKGSEASASRSQSDTSGPVTTTSSSISGFSYHLPSYKLVKLDVLAATHANYKVTLDLQ 1257
Db 1201 SPKKGSEASASRSQSDTSGPVTTTSSSISGFSYHLPSYKLVKLDVLAATHANYKVTLDLQ 1260
QY 1258 STEKFGGLRSALDVLQILELATLQIDICKVEETILGKSCFSREPMMAVTCVQOQLLKT 1317
Db 1261 STEKFGGLRSALDVLQILELATLQIDICKVEETILGKSCFSREPMMAVTCVQOQLLKT 1320
QY 1318 LFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1377

Db 1321 LFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFDVLRQVSTQKLTNTSVTKNRADKNAIHNHRLRFLPLVLIKALKQVTT 1437
Db 1381 VQAEQENDTSGWFDVLRQVSTQKLTNTSVTKNRADKNAIHNHRLRFLPLVLIKALKQVTT 1440
QY 1438 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDQVFIGVLYKQFEYIEVGQFRESEAIIPNIF 1497
Db 1441 TTCVOLQKQVLDLLAQLVQLRVNYCLLSDQVFIGVLYKQFEYIEVGQFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSHKQIIGIPKIIQICDGMASGRKA 1533
Db 1501 FFLVLLSYERYHSHKQIIGIPKIIQICDGMASGRKA 1536

RESULT 3
US-09-041-886-15
; Sequence 15 Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharoz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-886-15

Query Match 99.1%; Score 7821.5; DB 4; Length 3144;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLKSF---QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 57
Db 1 MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQA 60
QY 58 QPLLPPQPP 117
Db 61 QPLLPPQPP 120
QY 118 PEFQKLLGIAMELFLCSDDAESDVRVMADECLNKVTKALMDSNLPRLQLELYKEIKKNG 177
Db 121 PEFQKLLGIAMELFLCSDDAESDVRVMADECLNKVTKALMDSNLPRLQLELYKEIKKNG 180

QY 178 APRSLAALWRAELAHLPVKPCRYLVNLLPCLTRTSKRPEESVQETLAAPVKIMAS 237
DB 181 APRSLAALWRAELAHLPVKPCRYLVNLLPCLTRTSKRPEESVQETLAAPVKIMAS 240
QY 238 FGFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSCQHSRRRTQYFYSWLLNVLLGL 297
DB 241 FGFANDNEIKVLLKAFIANLKSSSTIRRTAAGSAVSCQHSRRRTQYFYSWLLNVLLGL 300
QY 298 LVPVEDEHSTLLTLGLVLLRLYLVLVLOOVKDTSLKSGFVTRKEMEVSPSAEOLVQVY 357
DB 301 LVPVEDEHSTLLTLGLVLLRLYLVLVLOOVKDTSLKSGFVTRKEMEVSPSAEOLVQVY 360
QY 358 ELTLHTHQHNDHNVTAAGLELQQLFRTPELLOLTAVGGTIGQLTAKEESGGRSRSG 417
DB 361 ELTLHTHQHNDHNVTAAGLELQQLFRTPELLOLTAVGGTIGQLTAKEESGGRSRSG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESROVSSSALTASVKDEISGEL 477
DB 421 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESROVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 537
DB 481 AASSGVSTPGSAGHDITTEQPSRQHTLQADSVDLASCDLTSSATDGDDEEDILSHSSQVS 540
QY 538 AVPSDAMDNDGQTQASSPISDSSTTTTEGPDSAVTPSDSSEIVLDGTDNQYIGLQIGQP 597
DB 541 AVPSDAMDNDGQTQASSPISDSSTTTTEGPDSAVTPSDSSEIVLDGTDNQYIGLQIGQP 600
QY 598 QDEDEBATGLPDSEAFRNSSMALQQAHLKKNMCHRCQPSDSSVDKFLRDEATEPGD 657
DB 601 QDEDEBATGLPDSEAFRNSSMALQQAHLKKNMCHRCQPSDSSVDKFLRDEATEPGD 660
QY 658 QENKPCRIKIDGSDDDSDAPLVHCVRLLSASFLLTGKGNVLVDPDRVRSVKALALSC 717
DB 661 QENKPCRIKIDGSDDDSDAPLVHCVRLLSASFLLTGKGNVLVDPDRVRSVKALALSC 720
QY 718 VGAVALHPESFFSKLYKVPDLTTEYPEQYVSDIILNYIDHGDQPVRGATAILCGFLICS 777
DB 721 VGAVALHPESFFSKLYKVPDLTTEYPEQYVSDIILNYIDHGDQPVRGATAILCGFLICS 780
QY 778 ILSRSPRHVGDMGTTIRLTGNTFSLADCIPLLRKTLKDESSVYTKLACTAVRNCVMSLC 837
DB 781 ILSRSPRHVGDMGTTIRLTGNTFSLADCIPLLRKTLKDESSVYTKLACTAVRNCVMSLC 840
QY 838 SSSVSELGLQIIDLTLRNSVWLVRTELETLAEIDFLVSLFLEAKAENLHRGAHHT 897
DB 841 SSSVSELGLQIIDLTLRNSVWLVRTELETLAEIDFLVSLFLEAKAENLHRGAHHT 900
QY 898 GLKLQERVLNNVYIHLGSDDEPRVRHVAASLIRLVKLFYKCDQGDQADPVVAVARDQS 957
DB 901 GLKLQERVLNNVYIHLGSDDEPRVRHVAASLIRLVKLFYKCDQGDQADPVVAVARDQS 960
QY 958 SVYLLKLMHETQPPSHFVSSTIIRIRGVNLLPSITDVTMNNLSRVIAAVSHELITSTT 1017
DB 961 SVYLLKLMHETQPPSHFVSSTIIRIRGVNLLPSITDVTMNNLSRVIAAVSHELITSTT 1020
QY 1018 RALTFCCEALCLLSTAFFVCINSGLWHCCVPPPLSASDESRSKCTVGMATMILTLLSSAW 1077
DB 1021 RALTFCCEALCLLSTAFFVCINSGLWHCCVPPPLSASDESRSKCTVGMATMILTLLSSAW 1080
QY 1078 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKOEVEWPALGDRALVPMV 1137
DB 1081 FPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEENPAATKOEVEWPALGDRALVPMV 1140
QY 1138 EQLFESHLKLVINICAHVLDVAVGPAKALPSLTNPSPSPIRKGEKEPEGEQASVPL 1197
DB 1141 EQLFESHLKLVINICAHVLDVAVGPAKALPSLTNPSPSPIRKGEKEPEGEQASVPL 1200
QY 1198 SPKKGSEASARSQSDTSGPVTTSSKSSLSGFYHLPSYKLHDVCLKATHANYKVTLDLQN 1257
DB 1201 SPKKGSEASARSQSDTSGPVTTSSKSSLSGFYHLPSYKLHDVCLKATHANYKVTLDLQN 1260
QY 1258 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLKSCFSREPMMATVCVQQLLKT 1317

DB 1261 STEKFGGFLRSALDVLSQLLELATLQDIGKCVBEILGYLKSCFSREPMMATVCVQQLLKT 1320
QY 1318 LFGNLSAQFQGLSSNPSKSGRAQLGSSVRPGLYHYCFMADYTHFTQALADASLRNM 1377
DB 1321 LFGNLSAQFQGLSSNPSKSGRAQLGSSVRPGLYHYCFMADYTHFTQALADASLRNM 1380
QY 1378 VQAEQENDTSGWFQVQKLVSTQKLTNTSVTKNRADKNAINHNHRLFEPLVIRKALKOYTT 1437
DB 1381 VQAEQENDTSGWFQVQKLVSTQKLTNTSVTKNRADKNAINHNHRLFEPLVIRKALKOYTT 1440
QY 1438 TTCVOLQKQVLDLALQVLQVRNYCLLSDSQVFTGVLKQFEYTEVQFRESEAIIPNIF 1497
DB 1441 TTCVOLQKQVLDLALQVLQVRNYCLLSDSQVFTGVLKQFEYTEVQFRESEAIIPNIF 1500
QY 1498 FFLVLLSYERYHSKOIIGIPKIIQICOGIMASGRKA 1533
DB 1501 FFLVLLSYERYHSKOIIGIPKIIQICOGIMASGRKA 1536

RESULT 4

US-08-246-982A-6
; Sequence 6, Application US/08246982A
; Patent No. 5686288
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Marcy E.
; APPLICANT: Ambrose, Christine M.
; APPLICANT: Duyao, Mabel P.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246.982A
; FILING DATE: May 20, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge, A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3880002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-246-982A-6

Query Match 98.9%; Score 7805.5; DB 1; Length 3144;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 1 MATLEKLMKAFESLSKSF---QQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQLPQPQA 57
DB 1 MATLEKLMKAFESLSKSFQQQQQQQQQQQQQQQQQQQQPPPPPPPPPPPPQLPQPQA 60
QY 58 QPLLPOQPP 117
DB 61 QPLLPOQPP 120

QY 118 PEFQKLLGIAEMLFLCSDDAESDVRMADECLNKVIALMDSNLRLOLELYKEIKKNG 177
DB 121 PEFQKLLGIAEMLFLCSDDAESDVRMADECLNKVIALMDSNLRLOLELYKEIKKNG 180
QY 178 APRSLAALWRFELAHVLPQCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKIMAS 237
DB 181 APRSLAALWRFELAHVLPQCRPYLVNLLPCLTRTSKRPEESVOETLAAAVPKIMAS 240
QY 238 FGFANDNEIKVALLKAFIANLKSSPTIRRTAAGSAVSIQHSRRTOYFYSWLLNVLGL 297
DB 241 FGFANDNEIKVALLKAFIANLKSSPTIRRTAAGSAVSIQHSRRTOYFYSWLLNVLGL 300
QY 298 LVPVEDEHSTLLILGVLLTLRYLVPLLOQQVKDTSLKSGFVTRKEMEVSPSAEQLVQVY 357
DB 301 LVPVEDEHSTLLILGVLLTLRYLVPLLOQQVKDTSLKSGFVTRKEMEVSPSAEQLVQVY 360
QY 358 ELTLHHTQHONVNVGALQLLQQLFRTPPPELLQTLTAVGGTGQJTAKEESGGRSRG 417
DB 361 ELTLHHTQHONVNVGALQLLQQLFRTPPPELLQTLTAVGGTGQJTAKEESGGRSRG 420
QY 418 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESRDVSYSALTASVKDEISGEL 477
DB 421 SIVELIAGGSSCPVLSRKQKGVLLGEEALEDDSESRDVSYSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDIITEPRSQHTLQADSVDLASCDLTSSATDGEDDILSHSSSQVS 537
DB 481 AASSGVSTPGSAGHDIITEPRSQHTLQADSVDLASCDLTSSATDGEDDILSHSSSQVS 540
QY 538 AVPSDPAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSSEIVLDGTDNQYLGLQIGQP 597
DB 541 AVPSDPAMDNDGTQASSPISDSQTTTEGPDSAVTPSDSSSEIVLDGTDNQYLGLQIGQP 600
QY 598 QDEDEATGLPDEASAFNSMAQQOAHLLKNMHCRCQPSDSSVDKFKVLRDEATEPGD 657
DB 601 QDEDEATGLPDEASAFNSMAQQOAHLLKNMHCRCQPSDSSVDKFKVLRDEATEPGD 660
QY 658 QENKPCRIKIDIGQSDDDSDAPLVHCVRLLSASFLLTGGKNVLPDRDVRVSVKALALSC 717
DB 661 QENKPCRIKIDIGQSDDDSDAPLVHCVRLLSASFLLTGGKNVLPDRDVRVSVKALALSC 720
QY 718 VGAVALHPESFSSKLYKVPDTEPEQYVSDILNYIDHGDQVQRGATAILCGTLICS 777
DB 721 VGAVALHPESFSSKLYKVPDTEPEQYVSDILNYIDHGDQVQRGATAILCGTLICS 780
QY 778 ILRSRPHVGDWGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 837
DB 781 ILRSRPHVGDWGTIRTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLC 840
QY 838 SSSYSELGLQIIDLVLTRNSSYWLVRTELLETFLAEIDFRLVFLSEAKAENLHRGAHHT 897
DB 841 SSSYSELGLQIIDLVLTRNSSYWLVRTELLETFLAEIDFRLVFLSEAKAENLHRGAHHT 900
QY 898 GLLKQBRVNNVVIHLLGDEDEPRVHVAASLIRLVKFLYKCDQGDQADPVVAVARDQS 957
DB 901 GLLKQBRVNNVVIHLLGDEDEPRVHVAASLIRLVKFLYKCDQGDQADPVVAVARDQS 960
QY 958 SVYKLLMHETQPPSHFSVTIRYRGYNLLPSITDVTMNNLSRVIAAVSHELITST 1017
DB 961 SVYKLLMHETQPPSHFSVTIRYRGYNLLPSITDVTMNNLSRVIAAVSHELITST 1020
QY 1018 RALTFCCEALCLLSTAFFVCIVSLGWHCVPPLSADSESRKSCVTGMMATMILTLLSSAW 1077
DB 1021 RALTFCCEALCLLSTAFFVCIVSLGWHCVPPLSADSESRKSCVTGMMATMILTLLSSAW 1080
QY 1078 FPLDLSAQDALILAGNLLAASAPKSLRSSWASEEPAANPAATKQEEVWPALGDALVPMV 1137
DB 1081 FPLDLSAQDALILAGNLLAASAPKSLRSSWASEEPAANPAATKQEEVWPALGDALVPMV 1140
QY 1138 EQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPPSLPRTRRKGKEPEGEQASVPL 1197
DB 1141 EQLFSLHLLKVINICAHVLDVAPGPAKALPSLTNPPSLPRTRRKGKEPEGEQASVPL 1200

QY 1198 SPKKGSEASAASRQSDTSGPVTTSSKSSSLGSPYHLPYSVLKLDVILKATHANYKVTLDLQN 1257
DB 1201 SPKKGSEASAASRQSDTSGPVTTSSKSSSLGSPYHLPYSVLKLDVILKATHANYKVTLDLQN 1260
QY 1258 STEKEGGFURSALDVLSQLLELATLQDICKVEEILGVLKSCFSREPMMATVCVQQLLKT 1317
DB 1261 STEKEGGFURSALDVLSQLLELATLQDICKVEEILGVLKSCFSREPMMATVCVQQLLKT 1320
QY 1318 LFGTNLASOFDGLSNPNKSGQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1377
DB 1321 LFGTNLASOFDGLSNPNKSGQRAQRLGSSSVRPGLYHYCFMAYTHFTQALADASLRNM 1380
QY 1378 VQAEQENTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNIHLRFLPELVIKALKQVTT 1437
DB 1381 VQAEQENTSGWFDVLQKVSTQKLTNLTSVTKNRADKNAIHNIHLRFLPELVIKALKQVTT 1440
QY 1438 TTCVOLQKQVLDLALQALVOLRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIF 1497
DB 1441 TTCVOLQKQVLDLALQALVOLRVNYCLLSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIF 1500
QY 1498 FFLVLLSVERYHSKQIIGIPKIIQICDGMASGRKA 1533
DB 1501 FFLVLLSVERYHSKQIIGIPKIIQICDGMASGRKA 1536

RESULT 5

US-08-453-265-6
; Sequence 6, Application US/08453265
; Patent No. 5693757
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Marcy E.
; APPLICANT: Ambrose, Christine M.
; APPLICANT: Duyao, Mabel P.
; APPLICANT: Gusella, James F.
; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,265
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3880003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-265-6

Query Match 98.9%; Score 7805.5; DB 1; Length 3144;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1530; Conservative 2; Mismatches 1; Indels 3; Gaps 1;
QY 1 MATLEKLMKAPSLKSF---QQQQQQQQQQQQQQQQQQPPPPPPPPPPPPPPPPQA 57
|||||

[illegible]

Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER OF SEQUENCES: 72	CORRESPONDENCE ADDRESS:	ADDRESSEE: Campbell & Flores LLP	STREET: 4370 La Jolla Village Drive, Suite 700	CITY: San Diego	STATE: California	COUNTRY: United States	ZIP: 92122	COMPUTER READABLE FORM:	MEDIUM TYPE: Floppy disk	COMPUTER: IBM PC compatible	OPERATING SYSTEM: PC-DOS/MS-DOS	SOFTWARE: Patent In Release #1.0, Version #1.25	CURRENT APPLICATION DATA:	APPLICATION NUMBER: US/09/041,886	FILING DATE:	CLASSIFICATION:	ATTORNEY/AGENT INFORMATION:	NAME: Campbell, Cathryn A.	REGISTRATION NUMBER: 31,815	REFERENCE/DOCKET NUMBER: P-LJ 2626	TELECOMMUNICATION INFORMATION:	TELEPHONE: (619) 535-9001	TELEFAX: (619) 535-8949	INFORMATION FOR SEQ ID NO: 29:	SEQUENCE CHARACTERISTICS:	LENGTH: 530 amino acids	TYPE: amino acid	TOPOLOGY: linear	MOLECULE TYPE: peptide	US-09-041-886-29
Db	QY	421	478	481	RESULT 12	US-09-041-886-28	Sequence 29, Application US/09041886	Patent No. 6235872	GENERAL INFORMATION:	APPLICANT: Bredesen, Dale E.	APPLICANT: Rabizadeh, Sharroz	TITLE OF INVENTION: Proopotic Peptides, Dependence	TITLE OF INVENTION: Polyptides and Methods of Use	NUMBER																														

QY 298 LVPVEDEHSTLLILGVLRLYLVPLQQVQKDTSLKSGFVTRKEMEVPSPAQLVQVY 357
Db 301 LVPVEDEHSTLLILGVLRLYLVPLQQVQKDTSLKSGFVTRKEMEVPSPAQLVQVY 360
QY 358 ELTLHTHQDHNVYTGALQLQLFRTPPPPELLOTLTAVGGIGQLTAKEESGGRSRSG 417
Db 361 ELTLHTHQDHNVYTGALQLQLFRTPPPPELLOTLTAVGGIGQLTAKEESGGRSRSG 420
QY 418 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESRSDVSSSALTASVKDEISGEL 477
Db 421 SIVELTAGGSSCPVLSRKQKGVLLGEEALEDDSESRSDVSSSALTASVKDEISGEL 480
QY 478 AASSGVSTPGSAGHDIIITEQPSQHTLQADSYD 510
Db 481 AASSGVSTPGSAGHDIIITEQPSQHTLQADSYD 513

RESULT 13

US-08-916-352-2
; Sequence 2, Application US/08916352
; Patent No. 6166191
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION
; TITLE OF INVENTION: HUMAN POLYHOMEOTIC 1 (hph1) ACTS AS A
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 HORTON STREET
; CITY: EMERYVILLE
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,352
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: POTTER, JANE
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 1355.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2707
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1004 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-916-352-2

Query Match 2.0%; Score 157.5; DB 4; Length 1004;
Best Local Similarity 39.1%; Pred. No. 0.002;
Matches 45; Conservative 8; Mismatches 33; Indels 29; Gaps 5;

QY 10 AFESLSFQQQQQ-----QQQQQQQQQQQQQQQ-----PPPPPPPPPPQL 50
Db 400 AIHQQQQFQHQSGQLLHTLQLAQQQQQQQQQQQQQQPPQATLTAPQPPVPTQV 459
QY 51 P-QPPQAQPLLPQ-----QPPPPPPPPPPGPAVAEPLHRPK-KELSATK 95
Db 460 PPSQSQQAQTLVQPMQLSSPLSLPPDAAPKPPPIQSKPPVAPIKPQLCAAK 514

RESULT 14

US-09-179-558-66

; Sequence 66, Application US/09179558
; Patent No. 6180612
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; APPLICANT: Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: TARGETING DNA METABOLIC PROCESSES USING
; TITLE OF INVENTION: AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/179,558
; FILING DATE: 27-OCT-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 09/060,470
; FILING DATE: 15-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 60/063,898
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 9426-005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEFAX: (212)8699741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-179-558-66

Query Match 2.0%; Score 156; DB 4; Length 311;
Best Local Similarity 23.8%; Pred. No. 0.0004;
Matches 91; Conservative 34; Mismatches 116; Indels 142; Gaps 20;

QY 38 PPPPPPPPPPPQLPQQPPQA-----QPLLPOQ-PPPPPPPPPPGPAVAEPLHRPK 88
Db 4 PPPPPPPPPPPPLPP 62
QY 89 KELSATKKDRVNHCLTICENIVAQSVRNSPEFQKLLGTAMELFLCSDDAESDVRVADE 148
Db 63 -ELFA-----SGALTPEPFKATSFQ--FPKSDQGGSPS---PDQ 98
QY 149 CLNKVIKALMDSNLPRLQLELYKEIKKNAPRSRLAALWRFALHLVRPKCRPYLVNL 208
Db 99 C-----QDQTPP-----GSPRRRX-----PPTK----- 118
QY 209 LPLTRTSKRPE-ESVQETLAAAVPKM-----ASFGNFANDNEIKVLLKAFIA 256
Db 119 ---LSSSKKKFSFSFNRTAEAKIPSTIEYILDLESGREKFLVFAHH---KVVLDAITK 172
QY 257 NLKSSSPITIR---TAAGSAVSICQHSRRRTQYFYSLNLNVL---LGLLVPEDEHSTLL 309
Db 173 ELEKRVQHIRIDGSTSSADRELCQQQLSPGPAVALSITANMGLIFSSAD----LV 228
QY 310 ILGLV-----LTLR-----YLVLQOQVQKDTSLKG-- 335

Query Match	2.08;	Score 155;	DB 1;	Length 2441;
Best Local Similarity	17.1%;	Pred. No. 0.013;		
Matches 128;	Conservative 93;	Mismatches 260;	Indels 266;	Gaps 26;
QY	13	SLKSFQQQQQQQQQQQQQQQQQ	-----	QQPPPPPPPPP-----PQLPQ 52
Db	1846	NKHNVRCQQIQHCLOAQLRRRMATNRNVPQSLPSFTSAPPETGTQOPSTPTQP 1905	::: :::	
QY	53	PPQAQPL-----	-----LPQPPPPPPP--	-PPCPAAVEEPLHR 86
Db	1906	PRAQPOPSPVNMSPAGFNNVARTOPTIVSAGKPTNPVAPPAEPQPPAAVEEAARQIER 1965	::: :::	
QY	87	PKKLSATKKDRVNHCL-----	-----TICENTVAOSVRNS-----	----- 117
Db	1966	EAOQQOHLRYANINMGMPGRDGMGTPCSQMTPVGLNVPRPNQVSGFVMSSMPPGOWQOA 2025	::: : :	
QY	118	--PEFOKKLLGTAMELFLCSDDAESDYRM-----	-----VADECLNKVIKALMDSNLPRLQL 167	
Db	2026	PIPQOOPGMPRPVMSMOAAAVAGPRMNPQNRSISPSALOQLLRLLTKSPSPQQOQ 2085	: : : : :	
QY	168	ELYKEIKKNCAPRSLRAALWFAELAHVLRQCRPYLVNLLPCLTRTSKRREESVQ----	224	
Db	2086	QVLNTLKSN---POLMAAFI-----	KORTAKYVAN-----QPQGQOPGLOSOP 2126	: : : : :

